

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:21:32 ; Search time 48 Seconds
(without alignments)
52.978 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLXENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 45 | 100.0 | 9 | 4 | AAB48919 | Generic S |
| 2 | 45 | 100.0 | 9 | 4 | AAB48917 | Aab48917 SH2 domai |
| 3 | 45 | 100.0 | 9 | 4 | AAB48922 | Aab48922 SH2 domai |
| 4 | 45 | 100.0 | 9 | 5 | ABG68582 | Abg68582 Peptide G |
| 5 | 45 | 100.0 | 10 | 4 | AAB48923 | Aab48923 SH2 domai |
| 6 | 45 | 100.0 | 10 | 4 | AAB48920 | Aab48920 SH2 domai |
| 7 | 45 | 100.0 | 10 | 4 | AAB48926 | Aab48926 SH2 domai |
| 8 | 45 | 100.0 | 10 | 4 | AAB48921 | Aab48921 SH2 domai |
| 9 | 45 | 100.0 | 10 | 4 | AAB48928 | Aab48928 SH2 domai |
| 10 | 45 | 100.0 | 11 | 2 | AAW46897 | AAW46897 GIC-S pep |
| 11 | 45 | 100.0 | 11 | 2 | AAW46896 | AAW46896 Non-phosp |
| 12 | 45 | 100.0 | 11 | 5 | ABG68419 | Abg68419 GI Peptid |
| 13 | 45 | 100.0 | 11 | 5 | ABG68583 | Abg68583 Peptide G |
| 14 | 45 | 100.0 | 26 | 4 | AAB48932 | Aab48932 SH2 domai |
| 15 | 45 | 100.0 | 26 | 4 | AAB48933 | Aab48933 SH2 domai |
| 16 | 37 | 82.2 | 11 | 2 | AAW46899 | AAW46899 Non-phosp |
| 17 | 37 | 82.2 | 919 | 2 | AAW63117 | AAW63117 Human ade |
| 18 | 36 | 80.0 | 11 | 2 | AAW46898 | AAW46898 Non-phosp |
| 19 | 36 | 80.0 | 20 | 2 | AAR49328 | Aar49328 Influenza |
| 20 | 36 | 80.0 | 20 | 2 | AAW54715 | AAW54715 Peptide f |
| 21 | 36 | 80.0 | 244 | 2 | AAW80804 | AAW80804 Amino aci |
| 22 | 36 | 80.0 | 244 | 2 | AAW95053 | AAW95053 Wyrotheci |
| 23 | 36 | 80.0 | 448 | 6 | ABU19327 | Abu19327 Protein e |
| 24 | 36 | 80.0 | 562 | 2 | AAR53588 | Aar53588 Full leng |
| 25 | 36 | 80.0 | 562 | 5 | AAE23111 | AAE23111 Influenza |

| | | | | | | |
|----|----|------|------|---|----------|--------------------|
| 26 | 36 | 80.0 | 921 | 6 | AAO23317 | Rhesus mo |
| 27 | 36 | 80.0 | 931 | 6 | AAO23313 | Aao23313 Cynomolgu |
| 28 | 35 | 77.8 | 9 | 2 | AAV10382 | T cell ep |
| 29 | 35 | 77.8 | 9 | 5 | ABG80064 | MHC class |
| 30 | 35 | 77.8 | 9 | 7 | ADC35620 | Influenza |
| 31 | 35 | 77.8 | 84 | 6 | ADA08462 | Human AFA |
| 32 | 35 | 77.8 | 86 | 6 | ADA08458 | Chicken A |
| 33 | 35 | 77.8 | 86 | 6 | ADA08461 | Avian AFA |
| 34 | 35 | 77.8 | 362 | 2 | AAV13465 | Peptide S |
| 35 | 35 | 77.8 | 634 | 4 | AAV13465 | Human pro |
| 36 | 35 | 77.8 | 815 | 6 | ADA08456 | Chicken A |
| 37 | 34 | 75.6 | 293 | 5 | ABG93283 | C. albica |
| 38 | 34 | 75.6 | 3542 | 4 | AAE62142 | P. falcip |
| 39 | 33 | 73.3 | 10 | 4 | AAE48925 | SH2 domai |
| 40 | 33 | 73.3 | 10 | 4 | AAE48927 | SH2 domai |
| 41 | 33 | 73.3 | 38 | 2 | AAE58364 | TSAR bind |
| 42 | 33 | 73.3 | 310 | 6 | ABM68832 | Phototrab |
| 43 | 33 | 73.3 | 434 | 4 | AAU33491 | Enterococ |
| 44 | 33 | 73.3 | 448 | 4 | AAU35058 | Enterococ |
| 45 | 33 | 73.3 | 448 | 6 | ABU14570 | Protein e |

ALIGNMENTS

RESULT 1
AAB48919
ID AAB48919 standard; peptide; 9 AA.

AC AAB48919;
XX
DT 16-MAR-2001 (first entry)
XX
DE Generic SH2 domain cyclic peptide inhibitor, SEQ ID NO:3.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytosolic; tumour; breast cancer; cyclic.
XX
OS Synthetic.

XX Key Location/Qualifiers
FT Modified-site 1..9
FT /note= "The nitrogen atoms of the N-terminus and the C-terminal amide are joined via a bridging moiety, thereby cyclising the peptide"
FT Misc-difference 1
FT /note= "Any naturally or non-naturally occurring amino acid except Glu"
FT Modified-site 9
FT /note= "C-terminal amide"

WO2000073326-A2.

07-DEC-2000.

02-JUN-2000; 2000WO-US015201.

02-JUN-1999; 99US-0137187P.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Roller PP, Long Y, Lung FT, King CR, Yang D;

WPI; 2001-137633/14.

Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src homology 2 domain binding to target protein, useful for preventing cancer, especially breast cancer.

Disclosure; Page 5; 26pp; English.

The invention relates to redox-stable, non-phosphorylated cyclic peptides which bind to Src homology 2 (SH2) domains, preventing them from binding

CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminoadipic acid (Aad), referred to as Adi in the specification; and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn reformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence is a generic
 CC representation of a cyclic peptide of the invention

XX Sequence 9 AA;

Query Match 100.0%; Score 45; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9

DB 1 XLYENVGMY 9

RESULT 2

AAB48917
 ID AAB48917 standard; peptide; 9 AA.

AC AAB48917;

DT 16-MAR-2001 (first entry)

DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:1.

XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
 KW cytosatic; tumour; breast cancer; cyclic.

OS Synthetic.

PH Key Location/Qualifiers

FT Modified-site 1.9
 /note= "The nitrogen atoms of the N-terminus and the C-
 FT terminal amide are joined via a bridging moiety, thereby
 FT cyclising the peptide"

FT Modified-site 1
 /note= "Gamma-carboxyglutamic acid"

FT Modified-site 9
 /note= "C-terminal amide"

XX WO200073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Roller PP, Long Y, Lung FT, King CR, Yang D;

XX WPI; 2001-137633/14.

XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 PT homology 2 domain binding to target protein, useful for preventing
 PT cancer, especially breast cancer.

XX

PS Claim 1; Page 21; 26pp; English.

XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to Src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminoadipic acid (Aad), referred to as Adi in the specification; and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn reformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a cyclic
 CC peptide of the invention

XX Sequence 9 AA;

Query Match 100.0%; Score 45; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9

DB 1 XLYENVGMY 9

RESULT 3

AAB48922
 ID AAB48922 standard; peptide; 9 AA.

AC AAB48922;

DT 16-MAR-2001 (first entry)

DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:7.

XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
 KW cytosatic; tumour; breast cancer; linear precursor.

OS Synthetic.

PH Key Location/Qualifiers

FT Modified-site 1
 /note= "Gamma-carboxyglutamic acid; the nitrogen atom of
 FT the N-terminus is joined to a ClCH2C(O) moiety"

FT Modified-site 9
 /note= "The carbon atom of the C-terminus is joined to a
 FT C(CH2SH)C(O)NH2 moiety"

XX WO200073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Roller PP, Long Y, Lung FT, King CR, Yang D;

XX WPI; 2001-137633/14.

PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 FT homology 2 domain binding to target protein, useful for preventing
 XX cancer, especially breast cancer.

PS Example 1; Page 13; 26pp; English.

XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to Src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a linear
 CC precursor of a peptide of the invention

XX
 SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMVY 9
 :|||||||
 DB 1 XLYENVGMVY 9

RESULT 4
 ID ABG68582 standard; peptide; 9 AA.
 XX
 AC ABG68582;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Peptide GITE #1.
 XX
 KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
 KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
 KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
 KW breast disorder; oesophageal disorder; pancreatic disorder; GI;
 KW prostate disorder; small intestine disorder; placental disorder;
 KW colon disorder; ovary disorder; testicular disorder; lung disorder.
 XX
 OS Synthetic.
 XX
 XX WO200236142-A2.
 PN
 XX 10-MAY-2002.
 PD
 XX 05-NOV-2001; 2001WO-US047400.
 PF
 XX 03-NOV-2000; 2000US-0245755P.
 PR
 XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 PA
 XX Krag DN, Pero SC, Oligino L;
 FI
 XX WPI; 2002-547451/58.
 DR
 XX Treatment or prophylaxis of a subject having a disorder characterized by

PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
 FT a non-phosphorylated peptide to a subject in need of the treatment.
 XX
 PS Disclosure; Fig 9B; 186pp; English.

XX The invention relates to treatment or prophylaxis (M1) of a subject
 CC having a disorder characterised by abnormal interaction of Grb7 (Growth
 CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
 CC administering to a subject in need of the treatment, a non-phosphorylated
 CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
 CC Asn) or its functional equivalent, in an amount effective to inhibit the
 CC disorder. Also included are peptide antagonists/inhibitors of Grb7;
 CC nucleic acids encoding the antagonists, an expression vector comprising
 CC the nucleic acid, a host cell transformed or transfected with the vector,
 CC screening (M2) a molecular library to identify a compound that inhibits
 CC interaction between Grb7 and a peptide antagonist and a phage display
 CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
 CC treatment of a subject having a disorder characterised by abnormal
 CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
 CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
 CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
 CC placenta, colon, ovary, testes and lung. The present sequence is a G1
 CC peptide (not defined) or derivative which is used to illustrate the
 CC possible structures of cyclic Grb7 antagonists

XX
 SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 5; Length 9;
 Best Local Similarity 88.9%; Pred. No. 1.4e+06;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMVY 9
 :|||||||
 DB 1 ELYENVGMVY 9

RESULT 5
 ID AAB48923 standard; peptide; 10 AA.
 XX
 AC AAB48923;
 XX
 DT 16-MAR-2001 (first entry)
 XX
 XX SH2 domain cyclic peptide inhibitor, SEQ ID NO:8.
 DE
 XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
 KW cytostatic; tumour; breast cancer; cyclic.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1..10
 FT /note= "The nitrogen atoms of the N-terminus and the C-
 FT terminal amide are joined via a bridging moiety, thereby
 FT cyclising the peptide"
 FT Modified-site 1
 FT /label= Aad
 FT Modified-site 10
 FT /note= "C-terminal amide"
 FT
 XX WO200073326-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX 02-JUN-2000; 2000WO-US015201.
 PF
 XX 02-JUN-1999; 99US-0137187P.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Roller PP, Long Y, Lung FT, King CR, Yang D;
 PI
 XX

DR WPI; 2001-137633/14.

XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.

XX Example 2; Page 13; 26pp; English.

XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention

XX Sequence 10 AA;

Query Match 100.0%; Score 45; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. NO. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLXENVGMVY 9
Db 1 XLXENVGMVY 9

RESULT 6

AAB48920
ID AAB48920 standard; peptide; 10 AA.

XX AAB48920;

XX 16-MAR-2001 (first entry)

XX SH2 domain cyclic peptide inhibitor, SEQ ID NO:4.

DE SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytostatic; tumour; breast cancer; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1.10
FT /note= "The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety C(O)-CH2-
FT S-CH2-CHC(O)NH2, thereby cyclising the peptide"

FT Modified-site 1
FT /note= "Gamma-carboxyglutamic acid"

FT Modified-site 10
FT /note= "C-terminal amide"

XX WO2000073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Rolier PP, Long Y, Lung FT, King CR, Yang D;

XX WPI; 2001-137633/14.

XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.

XX Example 1; Page 12; 26pp; English.

XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention

XX Sequence 10 AA;

Query Match 100.0%; Score 45; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. NO. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLXENVGMVY 9
Db 1 XLXENVGMVY 9

RESULT 7

AAB48926
ID AAB48926 standard; peptide; 10 AA.

XX AAB48926;

XX 16-MAR-2001 (first entry)

XX SH2 domain peptide inhibitor linear precursor, SEQ ID NO:11.

DE SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytostatic; tumour; breast cancer; linear precursor.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 10
FT /label= Nle
FT /note= "C-terminal amide, joined to a solid matrix"

XX WO2000073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

```

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX
XX PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
XX WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
XX PT homology 2 domain binding to target protein, useful for preventing
XX PT cancer, especially breast cancer.
XX
XX Example 4; Page 14; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
XX CC which bind to Src homology 2 (SH2) domains, preventing them from binding
XX CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
XX CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
XX CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
XX CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
XX CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
XX CC is either Aad or Glu. Optionally, there is a conservative or neutral
XX CC amino acid substitution at either or both of Leu2 and Gly7, and
XX CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
XX CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
XX CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
XX CC which links the nitrogen atom of the N terminus to the nitrogen atom of
XX CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
XX CC of less than 4.0 micromolar when the target protein is Grb2 (growth
XX CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
XX CC turn conformation. The peptides, and compositions comprising the
XX CC peptides, are useful for inhibiting the binding of the SH2 domain to a
XX CC target protein. They are particularly useful for preventing cancer,
XX CC especially breast cancer. The present sequence represents a linear
XX CC precursor of a peptide of the invention
XX
XX SQ Sequence 10 AA;
XX
XX Query Match 100.0%; Score 45; DB 4; Length 10;
XX Best Local Similarity 88.9%; Pred. No. 0.014;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 XLXENVGMVY 9
XX DB :|||||
XX 1 ELYENVGMVY 9
XX
XX RESULT 8
XX AAB48921
XX ID AAB48921 standard; peptide; 10 AA.
XX
XX AC AAB48921;
XX
XX DT 16-MAR-2001 (first entry)
XX
XX DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:5.
XX
XX KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX KW cytostatic; tumour; breast cancer; linear precursor.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT /note= "Gamma-carboxyglutamic acid"
XX
XX FN WO2000073326-A2.
XX
XX PD 07-DEC-2000.
XX
XX PF 02-JUN-2000; 2000WO-US015201.
XX
XX PR 02-JUN-1999; 99US-0137187P.
XX
XX

```

```

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
XX WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
XX PT homology 2 domain binding to target protein, useful for preventing
XX PT cancer, especially breast cancer.
XX
XX Example 1; Page 12; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
XX CC which bind to Src homology 2 (SH2) domains, preventing them from binding
XX CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
XX CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
XX CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
XX CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
XX CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
XX CC is either Aad or Glu. Optionally, there is a conservative or neutral
XX CC amino acid substitution at either or both of Leu2 and Gly7, and
XX CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
XX CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
XX CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
XX CC which links the nitrogen atom of the N terminus to the nitrogen atom of
XX CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
XX CC of less than 4.0 micromolar when the target protein is Grb2 (growth
XX CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
XX CC turn conformation. The peptides, and compositions comprising the
XX CC peptides, are useful for inhibiting the binding of the SH2 domain to a
XX CC target protein. They are particularly useful for preventing cancer,
XX CC especially breast cancer. The present sequence represents a linear
XX CC precursor of a peptide of the invention
XX
XX SQ Sequence 10 AA;
XX
XX Query Match 100.0%; Score 45; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.014;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 XLXENVGMVY 9
XX DB :|||||
XX 1 XLXENVGMVY 9
XX
XX RESULT 9
XX AAB48928
XX ID AAB48928 standard; peptide; 10 AA.
XX
XX AC AAB48928;
XX
XX DT 16-MAR-2001 (first entry)
XX
XX DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:14.
XX
XX KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX KW cytostatic; tumour; breast cancer; linear precursor.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Modified-site 10
XX FT /label= Aad
XX FT /note= "C-terminal amide, joined to a solid matrix"
XX
XX FN WO2000073326-A2.
XX
XX PD 07-DEC-2000.
XX
XX PF 02-JUN-2000; 2000WO-US015201.
XX
XX PR 02-JUN-1999; 99US-0137187P.
XX
XX

```

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX
 PI Roller PP, Long Y, Lung FT, King CR, Yang D;
 XX
 XX WPI; 2001-137633/14.
 XX
 XX
 PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the src
 PT homology 2 domain binding to target protein, useful for preventing
 PT cancer, especially breast cancer.
 XX
 XX Example 5; Page 15; 26pp; English.
 XX
 CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to Src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pYr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Glu); Xaa2 is 2-
 CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a linear
 CC precursor of a peptide of the invention
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 45; DB 4; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.014;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLXENVGMVY 9
 DB :|||||||
 1 ELXENVGMVY 9
 RESULT 10
 AAW46897
 ID AAW46897 standard; peptide; 11 AA.
 XX
 AC AAW46897;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE GIC-S peptide.
 XX
 KW SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
 KW signal transduction protein; non-phosphorylated; inhibition; treatment;
 KW hyper-proliferative disease; human cancer.
 XX
 OS Unidentified.
 XX
 XX WO9802176-A1.
 XX
 PD 22-JAN-1998.
 XX
 PF 16-JUL-1997; 97WO-US012501.
 XX
 PR 16-JUL-1996; 96US-0021858P.
 XX
 XX (GEOU) UNIV GEORGETOWN.
 XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX King CR, Sastry L, Krag D, Oligino L;
 XX
 PI

XX
 DR WPI; 1998-110340/10.
 XX
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein - at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.
 XX
 XX Disclosure; Page 18; 39pp; English.
 XX
 CC The present sequence represents a peptide designated GIC-S. This peptide
 CC is essentially the same as a non-phosphorylated peptide, G1, that is
 CC capable of binding to the src homology 2 (SH2) domain of Grb2, except
 CC that the terminal Cys residues of G1 are replaced with Ser residues. Grb2
 CC is a signal transduction protein. The binding affinity of the present
 CC peptide with Grb2 was tested, and it was demonstrated that the disulphide
 CC bond of G1 may be important. The G1 peptide binds to the SH2 domain of
 CC Grb2 with affinity similar to, or greater than, that of a SHC
 CC phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue
 CC that has not been modified by phosphate or similar charged group. The G1
 CC peptide is used to inhibit a signal transduction process that involves
 CC binding of a phosphorylated protein or peptide to the SH2 domain of a
 CC signal transduction protein, particularly Grb2. It is used specifically
 CC for treatment of hyper-proliferative diseases, especially human cancer
 XX
 SQ Sequence 11 AA;
 Query Match 100.0%; Score 45; DB 2; Length 11;
 Best Local Similarity 88.9%; Pred. No. 0.015;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLXENVGMVY 9
 DB :|||||||
 2 ELXENVGMVY 10
 RESULT 11
 AAW46896
 ID AAW46896 standard; peptide; 11 AA.
 XX
 AC AAW46896;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE Non-phosphorylated peptide which binds to the SH2 domain of Grb2.
 XX
 KW SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
 KW signal transduction protein; non-phosphorylated; inhibition; treatment;
 KW hyper-proliferative disease; human cancer; cyclic.
 XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FH Disulfide-bond 1. 11
 FT WO9802176-A1.
 XX
 PD 22-JAN-1998.
 XX
 PF 16-JUL-1997; 97WO-US012501.
 XX
 PR 16-JUL-1996; 96US-0021858P.
 XX
 XX (GEOU) UNIV GEORGETOWN.
 XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX King CR, Sastry L, Krag D, Oligino L;
 XX
 DR WPI; 1998-110340/10.
 XX
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein - at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.
 XX

PS Claim 9; Page 17; 39pp; English.

XX The present sequence represents non-phosphorylated peptide, G1, that is

CC capable of binding to the src homology 2 (SH2) domain of Grb2. Grb2 is a

CC signal transduction protein. The G1 peptide binds to the SH2 domain of

CC Grb2 with affinity similar to, or greater than, that of a SHC

CC phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue

CC that has not been modified by phosphate or similar charged group. The G1

CC peptide is used to inhibit a signal transduction process that involves

CC binding of a phosphorylated protein or peptide to the SH2 domain of a

CC signal transduction protein, particularly Grb2. It is used specifically

CC for treatment of hyper-proliferative diseases, especially human cancer

XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 45; DB 2; Length 11;

Best Local Similarity 88.9%; Pred. NO. 0.015;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9

DB 2 ELYENVGMY 10

RESULT 12

ABG68419

ID ABG68419 standard; peptide; 11 AA.

XX

AC ABG68419;

XX

DT 07-OCT-2002 (first entry)

XX

DE G1 peptide.

XX

KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;

KW cytosolic; cancer; phage display; tumour; metastasis; breast cancer;

KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;

KW breast disorder; oesophageal disorder; pancreatic disorder; G1;

KW prostate disorder; small intestine disorder; placental disorder;

KW colon disorder; ovary disorder; testicular disorder; lung disorder.

XX

OS Synthetic.

XX

PN WO200236142-A2.

XX

PD 10-MAY-2002.

XX

PF 05-NOV-2001; 2001WO-US047400.

XX

PR 03-NOV-2000; 2000US-0245755P.

XX

PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

XX

PI Krag DN, Pero SC, Oligino L;

XX

DR WPI; 2002-547451/58.

XX

PT Treatment or prophylaxis of a subject having a disorder characterized by

PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to

PT a non-phosphorylated peptide to a subject in need of the treatment.

XX

PS Disclosure; Page 102; 186pp; English.

XX

CC The invention relates to treatment or prophylaxis (M1) of a subject

CC having a disorder characterised by abnormal interaction of Grb7 (Growth

CC factor receptor-bound protein 7 and a Grb7 ligand, comprising

CC administering to a subject in need of the treatment, a non-phosphorylated

CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-

CC Asn) or its functional equivalent, in an amount effective to inhibit the

CC disorder. Also included are peptide antagonists/inhibitors of Grb7,

CC nucleic acids encoding the antagonists, an expression vector comprising

CC screening (M2) a molecular library to identify a compound that inhibits

CC interaction between Grb7 and a peptide antagonist and a phage display

CC the nucleic acid, a host cell transformed or transfected with the vector,

CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or

CC screening (M2) a molecular library to identify a compound that inhibits

CC interaction between Grb7 and a peptide antagonist and a phage display

CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or

CC treatment of a subject having a disorder characterised by abnormal

CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal

CC cancer, primary tumour or metastasis, or disorders in kidney, liver,

CC gonads, breast, oesophagus, pancreas, prostate, small intestine,

CC placenta, colon, ovary, testes and lung. The present sequence is a G1

CC peptide (not defined) or derivative which is used to illustrate the

CC possible structures of cyclic Grb7 antagonists

XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 45; DB 5; Length 11;

Best Local Similarity 88.9%; Pred. NO. 0.015;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9

DB 2 ELYENVGMY 10

RESULT 13

ABG68583

ID ABG68583 standard; peptide; 11 AA.

XX

AC ABG68583;

XX

DT 07-OCT-2002 (first entry)

XX

DE Peptide G1TE #2.

XX

KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;

KW cytosolic; cancer; phage display; tumour; metastasis; breast cancer;

KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;

KW breast disorder; oesophageal disorder; pancreatic disorder; G1;

KW prostate disorder; small intestine disorder; placental disorder;

KW colon disorder; ovary disorder; testicular disorder; lung disorder.

XX

OS Synthetic.

XX

PN WO200236142-A2.

XX

PD 10-MAY-2002.

XX

PF 05-NOV-2001; 2001WO-US047400.

XX

PR 03-NOV-2000; 2000US-0245755P.

XX

PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

XX

PI Krag DN, Pero SC, Oligino L;

XX

DR WPI; 2002-547451/58.

XX

PT Treatment or prophylaxis of a subject having a disorder characterized by

PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to

PT a non-phosphorylated peptide to a subject in need of the treatment.

XX

PS Disclosure; Fig 9C; 186pp; English.

XX

CC The invention relates to treatment or prophylaxis (M1) of a subject

CC having a disorder characterised by abnormal interaction of Grb7 (Growth

CC factor receptor-bound protein 7 and a Grb7 ligand, comprising

CC administering to a subject in need of the treatment, a non-phosphorylated

CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-

CC Asn) or its functional equivalent, in an amount effective to inhibit the

CC disorder. Also included are peptide antagonists/inhibitors of Grb7,

CC nucleic acids encoding the antagonists, an expression vector comprising

CC screening (M2) a molecular library to identify a compound that inhibits

CC interaction between Grb7 and a peptide antagonist and a phage display

CC the nucleic acid, a host cell transformed or transfected with the vector,

CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or

CC treatment of a subject having a disorder characterised by abnormal

CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
 CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
 CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
 CC placenta, colon, ovary testes and lung. The present sequence is a GI
 CC peptide (not defined) or derivative which is used to illustrate the
 CC possible structures of cyclic Grb7 antagonists
 XX
 SQ Sequence 11 AA;
 Query Match 100.0%; Score 45; DB 5; Length 11;
 Best Local Similarity 88.9%; Pred. No. 0.015;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVGVGY 9
 Db :|||||
 2 ELYENVGVGY 10
 RESULT 14
 AAB48932
 ID AAB48932 standard; peptide; 26 AA.
 XX
 AC AAB48932;
 XX
 DT 16-MAR-2001 (first entry)
 XX
 DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:18.
 XX
 KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
 KW cytosolic; tumour; breast cancer; linear precursor.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "Gamma-carboxyglutamic acid"
 FT
 XX
 XX WO200073326-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 02-JUN-2000; 2000WO-US015201.
 XX
 PR 02-JUN-1999; 99US-0137187P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Roller PP, Long Y, Lung FT, King CR, Yang D;
 XX
 DR WPI; 2001-137633/14.
 XX
 PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 PT homology 2 domain binding to target protein, useful for preventing
 PT cancer, especially breast cancer.
 XX
 PS Example 12; Page 19; 26pp; English.
 XX
 CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to Src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a

CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a linear
 CC precursor of a peptide of the invention
 XX
 SQ Sequence 26 AA;
 Query Match 100.0%; Score 45; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.042;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVGVGY 9
 Db :|||||
 1 XLYENVGVGY 9
 RESULT 15
 AAB48933
 ID AAB48933 standard; peptide; 26 AA.
 XX
 AC AAB48933;
 XX
 DT 16-MAR-2001 (first entry)
 XX
 DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:19.
 XX
 KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
 KW cytosolic; tumour; breast cancer; cyclic.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1.10
 FT /note= "The nitrogen atom of the N-terminus and the Cys
 FT 10 sidechain are joined via a bridging moiety, thereby
 FT cyclising part of the peptide"
 FT
 XX
 XX Modified-site 1 /note= "Gamma-carboxyglutamic acid"
 XX
 XX WO200073326-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 02-JUN-2000; 2000WO-US015201.
 XX
 PR 02-JUN-1999; 99US-0137187P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Roller PP, Long Y, Lung FT, King CR, Yang D;
 XX
 DR WPI; 2001-137633/14.
 XX
 PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 PT homology 2 domain binding to target protein, useful for preventing
 PT cancer, especially breast cancer.
 XX
 PS Example 12; Page 20; 26pp; English.
 XX
 CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to Src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a

CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention
XX

SQ Sequence 26 AA;

Query Match 100.0%; Score 45; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGY 9
| | | | | | | |
| | | | | | | |
DB 1 XLYENVGY 9

Search completed: July 15, 2004, 07:28:49
Job time : 51 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:26:37 ; Search time 14.5 Seconds
(without alignments)
32.044 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLYENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PCITUS_COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 39 | 86.7 | 566 | 2 | US-08-272-255-8 |
| 2 | 39 | 86.7 | 566 | 5 | PCT-US95-08565-8 |
| 3 | 37 | 82.2 | 919 | 2 | US-08-788-674-4 |
| 4 | 36 | 80.0 | 19 | 4 | US-09-376-343-3 |
| 5 | 36 | 80.0 | 20 | 2 | US-08-480-190-38 |
| 6 | 36 | 80.0 | 20 | 2 | US-08-488-379-38 |
| 7 | 36 | 80.0 | 20 | 4 | US-08-475-399A-38 |
| 8 | 36 | 80.0 | 20 | 5 | PCT-US93-07545-38 |
| 9 | 36 | 80.0 | 244 | 3 | US-09-003-287-6 |
| 10 | 36 | 80.0 | 244 | 3 | US-09-003-287-8 |
| 11 | 36 | 80.0 | 244 | 3 | US-09-518-988-2 |
| 12 | 35 | 77.8 | 9 | 1 | US-08-146-145-6 |
| 13 | 35 | 77.8 | 362 | 2 | US-09-080-897-6 |
| 14 | 35 | 77.8 | 362 | 3 | US-09-323-735-6 |
| 15 | 33 | 73.3 | 38 | 1 | US-08-176-500-22 |
| 16 | 33 | 73.3 | 38 | 1 | US-08-471-052A-22 |
| 17 | 33 | 73.3 | 38 | 1 | US-08-189-331-22 |
| 18 | 33 | 73.3 | 38 | 2 | US-08-471-939-22 |
| 19 | 33 | 73.3 | 38 | 2 | US-08-471-800-22 |
| 20 | 33 | 73.3 | 38 | 2 | US-08-471-068-22 |
| 21 | 33 | 73.3 | 245 | 4 | US-09-134-000C-3547 |
| 22 | 33 | 73.3 | 310 | 4 | US-09-252-991A-27339 |
| 23 | 33 | 73.3 | 693 | 4 | US-09-376-343-2 |
| 24 | 32 | 71.1 | 15 | 1 | US-08-176-500-31 |
| 25 | 32 | 71.1 | 15 | 1 | US-08-471-052A-31 |
| 26 | 32 | 71.1 | 15 | 1 | US-08-189-331-31 |
| 27 | 32 | 71.1 | 15 | 2 | US-08-471-939-31 |

| | | | | | | |
|----|----|------|-----|---|----------------------|-------------------|
| 28 | 32 | 71.1 | 15 | 2 | US-08-471-800-31 | Sequence 31, Appl |
| 29 | 32 | 71.1 | 15 | 2 | US-08-488-161-20 | Sequence 20, Appl |
| 30 | 32 | 71.1 | 15 | 2 | US-08-471-068-31 | Sequence 31, Appl |
| 31 | 32 | 71.1 | 15 | 3 | US-09-273-685-20 | Sequence 20, Appl |
| 32 | 32 | 71.1 | 15 | 5 | PCT-US95-11934-20 | Sequence 20, Appl |
| 33 | 32 | 71.1 | 445 | 4 | US-09-489-039A-13869 | Sequence 13869, A |
| 34 | 32 | 71.1 | 461 | 2 | US-08-527-227A-7 | Sequence 7, Appl |
| 35 | 32 | 71.1 | 485 | 4 | US-09-543-681A-4935 | Sequence 4935, Ap |
| 36 | 32 | 71.1 | 487 | 1 | US-08-249-112-4 | Sequence 4, Appl |
| 37 | 32 | 71.1 | 487 | 5 | PCT-US95-06556-4 | Sequence 4, Appl |
| 38 | 32 | 71.1 | 593 | 1 | US-08-202-389-12 | Sequence 12, Appl |
| 39 | 32 | 71.1 | 593 | 1 | US-08-018-129-5 | Sequence 5, Appl |
| 40 | 32 | 71.1 | 593 | 2 | US-08-448-250-5 | Sequence 5, Appl |
| 41 | 32 | 71.1 | 593 | 4 | US-09-282-257-5 | Sequence 5, Appl |
| 42 | 32 | 71.1 | 605 | 2 | US-08-752-307B-8 | Sequence 8, Appl |
| 43 | 32 | 71.1 | 605 | 4 | US-09-707-802-8 | Sequence 8, Appl |
| 44 | 32 | 71.1 | 605 | 4 | US-09-591-326-8 | Sequence 8, Appl |
| 45 | 32 | 71.1 | 671 | 3 | US-09-132-118-2 | Sequence 2, Appl |

ALIGNMENTS

RESULT 1
US-08-272-255-8
; Sequence 8, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLYING APPLICATION NUMBER: US/08/272,255
; FILING DATE: 08-JUL-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: UPN-1795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-272-255-8

Query Match 86.7%; Score 39; DB 2; Length 566;
Best Local Similarity 66.7%; Pred. No. 8.2;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 XLYENVGMY 9
Db 87 RLXNDVGLY 95

RESULT 2
PCT-US95-08565-8
; Sequence 8, Application PC/TUS9508565
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chenao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08565
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,255
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: UPN-1795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-08565-8

Query Match 86.7%; Score 39; DB 5; Length 566;
Best Local Similarity 66.7%; Pred. No. 8.2;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
:|||||:
Db 87 RLHNVGLY 95

RESULT 3
US-08-788-674-4
; Sequence 4, Application US/08788674
; Patent No. 5922315
; GENERAL INFORMATION:
; APPLICANT: Roy, Soumitra
; TITLE OF INVENTION: Adenoviruses Having Altered
; TITLE OF INVENTION: Hexon Proteins
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain,
; ADDRESSEE: Gilfillan, Cecchi, Stewart &
; ADDRESSEE: Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA

; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,674
; FILING DATE: 24-JAN-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 271010-363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 919 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: predicted hexon protein sequence
; NAME/KEY: for human Adenovirus 12
; US-08-788-674-4

Query Match 82.2%; Score 37; DB 2; Length 919;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
:|||||:
Db 439 FLYSNVGLY 447

RESULT 4
US-09-376-343-3
; Sequence 3, Application US/09376343
; Patent No. 6506592
; GENERAL INFORMATION:
; APPLICANT: Blum, Paul H.
; TITLE OF INVENTION: Hyperthermophilic Alpha-Glucosidase Gene and Its Use
; FILE REFERENCE: N1231-200
; CURRENT APPLICATION NUMBER: US/09/376,343
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/096,860
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
; US-09-376-343-3

Query Match 80.0%; Score 36; DB 4; Length 19;
Best Local Similarity 55.6%; Pred. No. 0.74;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
:|||||:
Db 5 KIYENLGVY 13

RESULT 5
US-08-480-190-38
; Sequence 38, Application US/08480190

```
; Patent No. 5827516
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,190
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-480-190-38
;
; Query Match 80.0%; Score 36; DB 2; Length 20;
; Best Local Similarity 66.7%; Pred. No. 0.78;
; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 XLYENVGY 9
; Db 2 TLYQNVGT 10
;
; RESULT 6
; US-08-488-379-38
; Sequence 38, Application US/08488379
; Patent No. 5880103
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
```

```
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,379
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-488-379-38
;
; Query Match 80.0%; Score 36; DB 2; Length 20;
; Best Local Similarity 66.7%; Pred. No. 0.78;
; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 XLYENVGY 9
; Db 2 TLYQNVGT 10
;
; RESULT 7
; US-08-475-399A-38
; Sequence 38, Application US/08475399A
; Patent No. 6509033
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicz, Roman M.
; APPLICANT: Vignali, Dario A. A.
; APPLICANT: Hedley, Mary L.
; APPLICANT: Stern, Lawrence J.
; APPLICANT: Strominger, Jack L.
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 276
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,399A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: 15-JUN-1993
```

APPLICATION NUMBER: 07/925,460
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00246/168003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-507
TELEFAX: 617/542-890
TELEX: 200154
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-475-399A-38

Query Match 80.0%; Score 36; DB 4; Length 20;
Best Local Similarity 66.7%; Pred. No. 0.78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMV 9
Db 2 TLYQNVGT 10

RESULT 8
PCT-US93-07545-38
GENERAL INFORMATION:
APPLICANT: Robert G. Urban
APPLICANT: Roman M. Chicz
APPLICANT: Dario A. A. Vignali
APPLICANT: Mary L. Hedley
APPLICANT: Lawrence J. Stern
APPLICANT: Jack L. Strominger
TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07545
FILING DATE: 19930811
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/925,460
FILING DATE: August 11, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00246/168001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 20
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US93-07545-38

Query Match 80.0%; Score 36; DB 5; Length 20;
Best Local Similarity 66.7%; Pred. No. 0.78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMV 9
Db 2 TLYQNVGT 10

RESULT 9
US-09-003-287-6
SEQUENCE 6, Application US/09003287
Patent No. 6096947
GENERAL INFORMATION:
APPLICANT: Jayne, Susan
APPLICANT: Barbour, Eric
APPLICANT: Meyer, Terry
FILE REFERENCE: mopat.mocah
CURRENT APPLICATION NUMBER: US/09/003,287
CURRENT FILING DATE: 1998-01-06
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 244
TYPE: PRT
ORGANISM: Myrothecium verrucaria
US-09-003-287-6

Query Match 80.0%; Score 36; DB 3; Length 244;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMV 9
Db 169 TLYDNVGV 177

RESULT 10
US-09-003-287-8
SEQUENCE 8, Application US/09003287
Patent No. 6096947
GENERAL INFORMATION:
APPLICANT: Jayne, Susan
APPLICANT: Barbour, Eric
APPLICANT: Meyer, Terry
FILE REFERENCE: mopat.mocah
CURRENT APPLICATION NUMBER: US/09/003,287
CURRENT FILING DATE: 1998-01-06
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 244
TYPE: PRT
ORGANISM: Myrothecium verrucaria
US-09-003-287-8

Query Match 80.0%; Score 36; DB 3; Length 244;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMV 9
Db 169 TLYDNVGV 177

RESULT 11
US-09-518-988-2
SEQUENCE 2, Application US/09518988
Patent No. 6268547
GENERAL INFORMATION:
APPLICANT: Weeks, James T.

;; TITLE OF INVENTION: TRANSFORMATION OF WHEAT WITH THE
;; TITLE OF INVENTION: CYANAMIDE HYDRATASE GENE
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Nancy J. Parsons
;; STREET: 800 Buchanan St.
;; CITY: Albany
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94710
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/518,988
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/873,001
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parsons, Nancy J.
;; REGISTRATION NUMBER: 40,364
;; REFERENCE/DOCKET NUMBER: 0177.95
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (510) 559-5731
;; TELEFAX: (510) 559-5736
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 244 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-518-988-2

Query Match 80.0%; Score 36; DB 3; Length 244;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 169 TLYDNVGY 177

RESULT 12
US-08-146-145-6
; Sequence 6, Application US/08-146145
; Patent No. 5747269
; GENERAL INFORMATION:
; APPLICANT: Rammensee, Hans-Georg
; APPLICANT: Falk, Kirsten
; APPLICANT: R tzsckke, Olaf
; APPLICANT: Stevanovic, Stefan
; APPLICANT: Jung, G nther
; TITLE OF INVENTION: DETERMINATION OF PEPTIDE MOTIFS ON MHC
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,145

;; FILING DATE: 17-NOV-1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kitts, Monica C.
;; REGISTRATION NUMBER: 36,105
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)638-5000
;; TELEFAX: (202)638-4810
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-146-145-6

Query Match 77.8%; Score 35; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
Db 1 LYQNVGTY 8

RESULT 13
US-09-080-897-6
; Sequence 6, Application US/09080897
; Patent No. 5985574
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan B.
; APPLICANT: Welcsh, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,897
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-080-897-6

Query Match 77.8%; Score 35; DB 2; Length 362;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
QY      1 XLYENVGMV 9
Db      247 KLYENLGEY 255

RESULT 14
US-09-323-735-6
; Sequence 6, Application US/09323735
; Patent No. 6197932
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan E.
; APPLICANT: Welch, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/323,735
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/080,897
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-323-735-6

Query Match      77.8%; Score 35; DB 3; Length 362;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 XLYENVGMV 9
Db      247 KLYENLGEY 255

RESULT 15
US-08-176-500-22
; Sequence 22, Application US/08176500
; Patent No. 5498538
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/013,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-143
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-176-500-22

Query Match      73.3%; Score 33; DB 1; Length 38;
Best Local Similarity 66.7%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 XLYENVGMV 9
Db      14 LLYANPGMY 22

Search completed: July 15, 2004, 07:31:18
Job time : 15.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:27:08 ; Search time 40 Seconds
(without alignments)
70.326 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLYENVGY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 45 | 100.0 | 9 | 10 | US-09-998-350-1 |
| 2 | 45 | 100.0 | 9 | 10 | US-09-998-350-3 |
| 3 | 45 | 100.0 | 9 | 10 | US-09-998-350-7 |
| 4 | 45 | 100.0 | 10 | 10 | US-09-998-350-4 |
| 5 | 45 | 100.0 | 10 | 10 | US-09-998-350-5 |
| 6 | 45 | 100.0 | 10 | 10 | US-09-998-350-6 |
| 7 | 45 | 100.0 | 10 | 10 | US-09-998-350-8 |
| 8 | 45 | 100.0 | 10 | 10 | US-09-998-350-11 |
| 9 | 45 | 100.0 | 10 | 10 | US-09-998-350-14 |
| 10 | 45 | 100.0 | 11 | 14 | US-10-013-815-32 |
| 11 | 45 | 100.0 | 26 | 10 | US-09-998-350-18 |
| 12 | 45 | 100.0 | 26 | 10 | US-09-998-350-19 |
| 13 | 36 | 80.0 | 244 | 15 | US-10-392-301-33 |
| 14 | 36 | 80.0 | 448 | 12 | US-10-282-122A-47251 |
| 15 | 35 | 77.8 | 9 | 12 | US-10-367-593-48 |

| | | | | | | |
|----|----|------|------|----|----------------------|--------------------|
| 16 | 35 | 77.8 | 9 | 12 | US-10-367-593-48 | Sequence 48, Appl |
| 17 | 35 | 77.8 | 9 | 12 | US-10-367-594-48 | Sequence 48, Appl |
| 18 | 35 | 77.8 | 9 | 12 | US-10-367-654-48 | Sequence 48, Appl |
| 19 | 35 | 77.8 | 9 | 12 | US-10-367-658-48 | Sequence 48, Appl |
| 20 | 35 | 77.8 | 9 | 12 | US-10-367-668-48 | Sequence 48, Appl |
| 21 | 35 | 77.8 | 9 | 16 | US-10-367-674-48 | Sequence 48, Appl |
| 22 | 35 | 77.8 | 9 | 16 | US-10-777-053-366 | Sequence 366, App |
| 23 | 35 | 77.8 | 9 | 16 | US-10-777-053-943 | Sequence 943, App |
| 24 | 35 | 77.8 | 9 | 16 | US-10-777-053-958 | Sequence 958, App |
| 25 | 35 | 77.8 | 79 | 14 | US-10-246-354-7 | Sequence 7, Appl |
| 26 | 35 | 77.8 | 84 | 14 | US-10-246-354-10 | Sequence 10, Appl |
| 27 | 35 | 77.8 | 86 | 14 | US-10-246-354-6 | Sequence 6, Appl |
| 28 | 35 | 77.8 | 168 | 12 | US-10-424-599-170035 | Sequence 170035, A |
| 29 | 35 | 77.8 | 815 | 14 | US-10-246-354-3 | Sequence 3, Appl |
| 30 | 35 | 77.8 | 1096 | 16 | US-10-408-765A-747 | Sequence 747, App |
| 31 | 34 | 75.6 | 1234 | 15 | US-10-369-493-13287 | Sequence 13287, A |
| 32 | 34 | 75.6 | 3542 | 12 | US-10-087-013-2 | Sequence 2, Appl |
| 33 | 33 | 73.3 | 73.3 | 10 | US-09-998-350-10 | Sequence 10, Appl |
| 34 | 33 | 73.3 | 73.3 | 10 | US-09-998-350-12 | Sequence 12, Appl |
| 35 | 33 | 73.3 | 73.3 | 10 | US-09-998-350-13 | Sequence 13, Appl |
| 36 | 33 | 73.3 | 73.3 | 78 | US-10-424-599-219681 | Sequence 219681, A |
| 37 | 33 | 73.3 | 134 | 16 | US-10-437-963-168439 | Sequence 168439, A |
| 38 | 33 | 73.3 | 162 | 12 | US-10-424-599-205104 | Sequence 205104, A |
| 39 | 33 | 73.3 | 306 | 15 | US-10-369-493-1088 | Sequence 1088, Ap |
| 40 | 33 | 73.3 | 434 | 9 | US-09-815-242-4987 | Sequence 4987, Ap |
| 41 | 33 | 73.3 | 448 | 9 | US-09-815-242-10651 | Sequence 10651, A |
| 42 | 33 | 73.3 | 448 | 12 | US-10-282-122A-42494 | Sequence 42494, A |
| 43 | 33 | 73.3 | 693 | 14 | US-10-228-063-5 | Sequence 5, Appl |
| 44 | 33 | 73.3 | 712 | 14 | US-10-228-063-27 | Sequence 27, Appl |
| 45 | 33 | 73.3 | 718 | 14 | US-10-228-063-26 | Sequence 26, Appl |

ALIGNMENTS

RESULT 1

US-09-998-350-1
; Sequence 1, Application US/0998350
; Publication No US20030079368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Datun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is an amide, i.e. C(O)NH
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (1)..(9)
 ; OTHER INFORMATION: Xaa (Gla) and Tyr at position 9 are bridged together, making this
 ; OTHER INFORMATION: peptide cyclic
 US-09-998-350-1

Query Match 100.0%; Score 45; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
 |||||
 Db 1 XLYENVGMVY 9

RESULT 2

US-09-998-350-3
 ; Sequence 3, Application US/09998350
 ; Publication No. US20030078368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roller, Peter P
 ; APPLICANT: Long, Ya-Qiu
 ; APPLICANT: Lung, Feng-Di T
 ; APPLICANT: King, Richter C
 ; APPLICANT: Yang, Dajun
 ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 ; FILE OF INVENTION: SYNTHESIS AND USE
 ; FILE REFERENCE: 214683
 ; CURRENT APPLICATION NUMBER: US/09/998,350
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: PCT/US00/15201
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 60/137,187
 ; PRIOR FILING DATE: 1999-06-02
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1)
 ; FEATURE:
 ; OTHER INFORMATION: Xaa is any amino acid other than Glu
 ; NAME/KEY: misc_feature
 ; LOCATION: (9)..(9)
 ; FEATURE:
 ; OTHER INFORMATION: Tyr at position 9 is an amide, i.e., C(O)NH

Query Match 100.0%; Score 45; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
 |||||
 Db 1 XLYENVGMVY 9

RESULT 3

US-09-998-350-7
 ; Sequence 7, Application US/09998350
 ; Publication No. US20030078368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roller, Peter P
 ; APPLICANT: Long, Ya-Qiu
 ; APPLICANT: Lung, Feng-Di T
 ; APPLICANT: King, Richter C
 ; APPLICANT: Yang, Dajun
 ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 ; FILE OF INVENTION: SYNTHESIS AND USE
 ; FILE REFERENCE: 214683
 ; CURRENT APPLICATION NUMBER: US/09/998,350
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: PCT/US00/15201
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 60/137,187
 ; PRIOR FILING DATE: 1999-06-02
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1)
 ; FEATURE:
 ; OTHER INFORMATION: Xaa is any amino acid other than Glu
 ; NAME/KEY: misc_feature
 ; LOCATION: (9)..(9)
 ; FEATURE:
 ; OTHER INFORMATION: Tyr at position 9 is an amide, i.e., C(O)NH

Query Match 100.0%; Score 45; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
 |||||
 Db 1 XLYENVGMVY 9

RESULT 4

US-09-998-350-4
 ; Sequence 4, Application US/09998350
 ; Publication No. US20030078368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roller, Peter P
 ; APPLICANT: Long, Ya-Qiu
 ; APPLICANT: Lung, Feng-Di T
 ; APPLICANT: King, Richter C
 ; APPLICANT: Yang, Dajun
 ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 ; FILE OF INVENTION: SYNTHESIS AND USE
 ; FILE REFERENCE: 214683
 ; CURRENT APPLICATION NUMBER: US/09/998,350
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: PCT/US00/15201
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 60/137,187
 ; PRIOR FILING DATE: 1999-06-02
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1)
 ; FEATURE:
 ; OTHER INFORMATION: Xaa and Tyr at position 9 are bridged together, making this peptide cyclic
 ; NAME/KEY: misc_feature
 ; LOCATION: (9)..(9)
 ; FEATURE:
 ; OTHER INFORMATION: de cyclic

Query Match 100.0%; Score 45; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
 |||||
 Db 1 XLYENVGMVY 9

APPLICANT: Lung, Feng-Di T
 APPLICANT: King, Richter C
 APPLICANT: Yang, Dajun
 TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 FILE OF INVENTION: SYNTHESIS AND USE
 FILE REFERENCE: 214683
 CURRENT APPLICATION NUMBER: US/09/998,350
 CURRENT FILING DATE: 2002-12-09
 PRIOR APPLICATION NUMBER: PCT/US00/15201
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: 60/137,187
 PRIOR FILING DATE: 1999-06-02
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 7
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(1)
 OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(1)
 OTHER INFORMATION: Xaa has a ClCH2C(O)- group attached
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (9)..(9)
 OTHER INFORMATION: Tyr at position 9 has a -C(CH2SH)C(O)NH2 group attached
 US-09-998-350-7

Query Match 100.0%; Score 45; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
 |||||
 Db 1 XLYENVGMVY 9

RESULT 4

US-09-998-350-4
 ; Sequence 4, Application US/09998350
 ; Publication No. US20030078368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roller, Peter P
 ; APPLICANT: Long, Ya-Qiu
 ; APPLICANT: Lung, Feng-Di T
 ; APPLICANT: King, Richter C
 ; APPLICANT: Yang, Dajun
 ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 ; FILE OF INVENTION: SYNTHESIS AND USE
 ; FILE REFERENCE: 214683
 ; CURRENT APPLICATION NUMBER: US/09/998,350
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: PCT/US00/15201
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 60/137,187
 ; PRIOR FILING DATE: 1999-06-02
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1)
 ; FEATURE:
 ; OTHER INFORMATION: Xaa and Tyr at position 9 are bridged together, making this peptide cyclic
 ; NAME/KEY: misc_feature
 ; LOCATION: (9)..(9)
 ; FEATURE:
 ; OTHER INFORMATION: de cyclic

Query Match 100.0%; Score 45; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMVY 9
 |||||
 Db 1 XLYENVGMVY 9

```

; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys at position 10 is an amide, i.e., C(O)NH
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Xaa (Gla) and Cys are bridged together, making this peptide cyclic
; OTHER INFORMATION: C
; US-09-998-350-4

```

```

Query Match          100.0%; Score 45; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 XLYENVGMV 9
    |||||
Db 1 XLYENVGMV 9

```

```

RESULT 5
US-09-998-350-5
; Sequence 5, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
; US-09-998-350-5

```

```

Query Match          100.0%; Score 45; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 XLYENVGMV 9
    |||||
Db 1 XLYENVGMV 9

```

```

RESULT 6
US-09-998-350-6
; Sequence 6, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu

```

```

; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND N
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla(OtBu)2, which is di-tert-butoxy-gamma-carboxy-L-glutam
; OTHER INFORMATION: ic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Tyr at position 3 is modified to Tyr(tBu), which is tert-butyl-ty
; OTHER INFORMATION: rosine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is is trytyl-asp
; OTHER INFORMATION: argine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(tBu), which is tert-butyl-ty
; OTHER INFORMATION: rosine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys at position 10 is modified to Cys(Trt), which is trytyl-cyste
; OTHER INFORMATION: ine, and Cys(Trt) is connected to a resin
; US-09-998-350-6

```

```

Query Match          100.0%; Score 45; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 XLYENVGMV 9
    |||||
Db 1 XLYENVGMV 9

```

```

RESULT 7
US-09-998-350-8
; Sequence 8, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2

```

| | |
|---|--|
| <pre> ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND ; TITLE OF INVENTION: SYNTHESIS AND USE ; FILE REFERENCE: 214683 ; CURRENT APPLICATION NUMBER: US/09/998,350 ; CURRENT FILING DATE: 2002-12-09 ; PRIOR APPLICATION NUMBER: PCT/US00/15201 ; PRIOR FILING DATE: 2000-06-02 ; PRIOR APPLICATION NUMBER: 60/137,187 ; PRIOR FILING DATE: 1999-06-02 ; NUMBER OF SEQ ID NOS: 19 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 8 ; LENGTH: 10 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Synthetic ; NAME/KEY: misc feature ; LOCATION: (1)..(1) ; OTHER INFORMATION: Xaa = Adi, which is alpha-amino-adipic acid ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (1)..(1) ; OTHER INFORMATION: Xaa has a CH2CO- group attached ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (1)..(1) ; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (1)..(1) ; OTHER INFORMATION: Cys are bridged together, making this peptide cycli US-09-998-350-8 </pre> | <pre> ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND ; TITLE OF INVENTION: SYNTHESIS AND USE ; FILE REFERENCE: 214683 ; CURRENT APPLICATION NUMBER: US/09/998,350 ; CURRENT FILING DATE: 2002-12-09 ; PRIOR APPLICATION NUMBER: PCT/US00/15201 ; PRIOR FILING DATE: 2000-06-02 ; PRIOR APPLICATION NUMBER: 60/137,187 ; PRIOR FILING DATE: 1999-06-02 ; NUMBER OF SEQ ID NOS: 19 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 8 ; LENGTH: 10 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Synthetic ; NAME/KEY: misc feature ; LOCATION: (1)..(1) ; OTHER INFORMATION: glutamic acid ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (3)..(3) ; OTHER INFORMATION: Tyr at position 3 is modified to Tyr(OtBu), which is tert-butoxy- ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (4)..(4) ; OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy- ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (5)..(5) ; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trityl-aspara ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (9)..(9) ; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy- ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (10)..(10) ; OTHER INFORMATION: Xaa = Nle, which is norleucine ; FEATURE: ; NAME/KEY: misc feature ; LOCATION: (10)..(10) ; OTHER INFORMATION: Xaa is an amide and is attached to a resin US-09-998-350-11 </pre> |
| <pre> Query Match 100.0%; Score 45; DB 10; Length 10; Best Local Similarity 100.0%; Pred. NO. 0.034; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 1 XLVENVGMY 9 Db 1 XLVENVGMY 9 </pre> | <pre> Query Match 100.0%; Score 45; DB 10; Length 10; Best Local Similarity 88.9%; Pred. NO. 0.034; Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0; QY 1 XLVENVGMY 9 : Db 1 ELVENVGMY 9 </pre> |
| <pre> RESULT 8 US-09-998-350-11 ; Sequence 11, Application US/09998350 ; Publication No. US20030078368A1 ; GENERAL INFORMATION: ; APPLICANT: Roller, Peter P ; APPLICANT: Long, Ya-Qiu ; APPLICANT: Lung, Feng-Di T ; APPLICANT: King, Richter C ; APPLICANT: Yang, Dajun ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND ; TITLE OF INVENTION: SYNTHESIS AND USE ; FILE REFERENCE: 214683 ; CURRENT APPLICATION NUMBER: US/09/998,350 ; CURRENT FILING DATE: 2002-12-09 ; PRIOR APPLICATION NUMBER: PCT/US00/15201 ; PRIOR FILING DATE: 2000-06-02 ; PRIOR APPLICATION NUMBER: 60/137,187 ; PRIOR FILING DATE: 1999-06-02 ; NUMBER OF SEQ ID NOS: 19 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 11 ; LENGTH: 10 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Synthetic </pre> | <pre> RESULT 9 US-09-998-350-14 ; Sequence 14, Application US/09998350 ; Publication No. US20030078368A1 ; GENERAL INFORMATION: ; APPLICANT: Roller, Peter P ; APPLICANT: Long, Ya-Qiu ; APPLICANT: Lung, Feng-Di T ; APPLICANT: King, Richter C ; APPLICANT: Yang, Dajun ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND ; TITLE OF INVENTION: SYNTHESIS AND USE ; FILE REFERENCE: 214683 ; CURRENT APPLICATION NUMBER: US/09/998,350 ; CURRENT FILING DATE: 2002-12-09 ; PRIOR APPLICATION NUMBER: PCT/US00/15201 ; PRIOR FILING DATE: 2000-06-02 ; PRIOR APPLICATION NUMBER: 60/137,187 ; PRIOR FILING DATE: 1999-06-02 ; NUMBER OF SEQ ID NOS: 19 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 14 ; LENGTH: 10 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Synthetic </pre> |

```
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Glu at position 1 is modified to Glu(OtBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trytyl-aspara-
; OTHER INFORMATION: gine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy-
; OTHER INFORMATION: tyrosine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa = Adi(OAl), which is allyloxy-alpha-amino-adipic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa is an amide, i.e., C(O)NH
US-09-998-350-14
```

```
Query Match 100.0%; Score 45; DB 10; Length 10;
Best Local Similarity 88.9%; Pred.No. 0.034; 0; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0;
```

```
QY 1 XLYENVGMVY 9
Db 1 ELYENVGMVY 9
```

RESULT 10

```
US-10-013-815-32
; Sequence 32, Application US/10013815
; Publication No. US20030105000A1
; GENERAL INFORMATION:
; APPLICANT: Pero, Stephanie
; APPLICANT: Krag, David
; APPLICANT: Oligino, Lyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING GRB7
; FILE REFERENCE: V0139/7048 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/10/013,815
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 60/245,755
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030105000A1-phosphorylated peptide with YEN motif
US-10-013-815-32
```

```
Query Match 100.0%; Score 45; DB 14; Length 11;
Best Local Similarity 88.9%; Pred.No. 0.037; 0; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0;
```

```
QY 1 XLYENVGMVY 9
Db 2 ELYENVGMVY 10
```

RESULT 11

```
US-09-998-350-18
```

```
; Sequence 18, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND USE
; FILE REFERENCES: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
US-09-998-350-18
```

```
Query Match 100.0%; Score 45; DB 10; Length 26;
Best Local Similarity 100.0%; Pred.No. 0.095; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;
```

```
QY 1 XLYENVGMVY 9
Db 1 XLYENVGMVY 9
```

RESULT 12

```
US-09-998-350-19
; Sequence 19, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
```

FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(1)
OTHER INFORMATION: Xaa (Gla) has a CH2CO- group attached
FEATURE:
NAME/KEY: misc feature
LOCATION: (10)-(10)
OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
US-09-998-350-19

Query Match 100.0%; Score 45; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLXENVGMV 9
| | | | | | | | | |
Db 1 XLXENVGMV 9

RESULT 13
US-10-392-301-33
Sequence 33, Application US/10392301
Publication No. US20040003434A1
GENERAL INFORMATION:
APPLICANT: WEEKS, J. TROY
APPLICANT: ROMWENS, CAIUS
TITLE OF INVENTION: REFINED PLANT TRANSFORMATION
FILE REFERENCE: 058951/0164
CURRENT APPLICATION NUMBER: US/10/392,301
CURRENT FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: 60/365,527
PRIOR FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: 60/377,597
PRIOR FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 33
LENGTH: 244
TYPE: PRT
ORGANISM: Myrothecium verrucaria
US-10-392-301-33

Query Match 80.0%; Score 36; DB 15; Length 244;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLXENVGMV 9
: | | | | | | | | | |
Db 169 TLYDNVGV 177

RESULT 14
US-10-282-122A-47251
Sequence 47251, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47251
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-282-122A-47251

Query Match 80.0%; Score 36; DB 12; Length 448;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLXENVGMV 9
: | | | | | | | | | |
Db 336 LLYEDIGLY 344

RESULT 15
US-10-367-580-48
Sequence 48, Application US/10367580
Publication No. US20040071720A1
GENERAL INFORMATION:
APPLICANT: Rothman, James E.
APPLICANT: Hartl, F. Ulrich
APPLICANT: Hoe, Mee H.
APPLICANT: Houghton, Alan
APPLICANT: Takechi, Yoshizumi
APPLICANT: Mayhew, Mark
TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
FILE REFERENCE: 11746/461061
CURRENT APPLICATION NUMBER: US/10/367,580
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US 09/794,832
PRIOR FILING DATE: 2001-02-27
PRIOR APPLICATION NUMBER: US 09/011,645
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: PCT/US96/13363
PRIOR FILING DATE: 1996-08-16
PRIOR APPLICATION NUMBER: US 60/002,490
PRIOR FILING DATE: 1995-08-18
PRIOR APPLICATION NUMBER: US 60/002,479
PRIOR FILING DATE: 1995-08-18
NUMBER OF SEQ ID NOS: 349
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 48
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-10-367-580-48

Query Match 77.8%; Score 35; DB 12; Length 9;

Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 2 LYENVGY 9
 ||:||||
Db 1 LYQNVGY 8

Search completed: July 15, 2004, 07:32:49
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:23:22 ; Search time 11.5 Seconds
(without alignments)
75.280 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLYENVGMV 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|------------------------|
| 1 | 39 | 86.7 | 565 | 2 S67298 | deoxyribodipyrimid |
| 2 | 37 | 82.2 | 468 | 2 S37217 | hexon protein - hu |
| 3 | 37 | 82.2 | 526 | 1 VGN5G | spike glycoprotein |
| 4 | 37 | 82.2 | 919 | 2 S33942 | hexon protein - hu |
| 5 | 36 | 80.0 | 20 | 2 P0161 | hemagglutinin - in |
| 6 | 36 | 80.0 | 244 | 2 A39365 | cyanamide hydratase |
| 7 | 36 | 80.0 | 447 | 2 S32956 | hexon protein - hu |
| 8 | 36 | 80.0 | 448 | 1 F70190 | probable diphosphatase |
| 9 | 36 | 80.0 | 562 | 1 HMTV2 | hemagglutinin prec |
| 10 | 36 | 80.0 | 936 | 2 S57637 | hexon protein - hu |
| 11 | 35 | 77.8 | 29 | 2 B81136 | hypothetical prote |
| 12 | 35 | 77.8 | 34 | 2 H81883 | hypothetical prote |
| 13 | 35 | 77.8 | 150 | 2 A55883 | actin-filament-ass |
| 14 | 34 | 75.6 | 511 | 2 A99574 | ABC transporter at |
| 15 | 34 | 75.6 | 1249 | 2 A56511 | myosin I myoA - Em |
| 16 | 33 | 73.3 | 99 | 2 S44632 | f22b7.3 protein - |
| 17 | 33 | 73.3 | 306 | 2 D64497 | aspartate carbanoy |
| 18 | 33 | 73.3 | 309 | 2 F83044 | nitrate-inducible |
| 19 | 33 | 73.3 | 312 | 2 A64042 | formate dehydrogen |
| 20 | 33 | 73.3 | 313 | 2 A57499 | N5-(carboxyethyl)o |
| 21 | 33 | 73.3 | 332 | 2 T33774 | hypothetical prote |
| 22 | 33 | 73.3 | 352 | 2 D72264 | hypothetical prote |
| 23 | 33 | 73.3 | 354 | 2 E97128 | magnesium and ciba |
| 24 | 33 | 73.3 | 389 | 2 B81380 | hypothetical prote |
| 25 | 33 | 73.3 | 439 | 2 G88103 | protein M10G11.17 |
| 26 | 33 | 73.3 | 512 | 1 VGIIVH | envelope glycoprot |
| 27 | 33 | 73.3 | 661 | 2 S49901 | coat protein gpi - |
| 28 | 33 | 73.3 | 688 | 2 T33708 | hypothetical prote |
| 29 | 33 | 73.3 | 693 | 2 H90486 | alpha-glucosidase |

30 33 73.3 700 2 T20550 hypothetical prote
31 33 73.3 739 2 A11876 hypothetical prote
32 33 73.3 852 2 T33824 hypothetical prote
33 32 71.1 149 2 S67188 hypothetical prote
34 32 71.1 221 2 E84400 conserved hypothet
35 32 71.1 224 2 H89847 conserved hypothet
36 32 71.1 231 2 H85138 conserved hypothet
37 32 71.1 263 2 A10471 hypothetical prote
38 32 71.1 307 2 T34973 probable exported
39 32 71.1 336 2 T30459 5,10-methylenetet
40 32 71.1 434 2 S50865 hypothetical prote
41 32 71.1 437 2 A64891 avermectin-sensiti
42 32 71.1 450 1 DCCRO coenzyme F390 synt
43 32 71.1 455 1 DCHYOC ornithine decarbox
44 32 71.1 460 2 A43563 ornithine decarbox
45 32 71.1 461 1 DCHUO ornithine decarbox

ALIGNMENTS

RESULT 1

S67298

deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O6771; protein YOR386w
C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
C:Accession: S67298; A23964; A24046

R:Delius, H.; Hebling, U.; Hofmann, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67261

A:Accession: S67298

A:Molecule type: DNA

A:Residues: 1-565

A:Cross-references: EMBL:Z75294; NID:G1420830; PIDN:CAA99718.1; PID:G1420831; MIPS:YOR386

A:Experimental source: strain S288C

R:Yasui, A.; Langeveid, S.A.

Gene 36, 349-355, 1985

A:Title: Homology between the photoreactivation genes of Saccharomyces cerevisiae and Esc
A:Reference number: A23964; MUID:86083177; PMID:3000886

A:Accession: A23964

A:Molecule type: DNA

A:Residues: 1-76, 'A', '78-164, 'S', '166-168, 'T', '170-199, 'S', '201-350, 'R', '352-364, 'E', '366-472, '
A:Cross-references: EMBL:M11578; NID:G172169; PIDN:AAA34875.1; PID:G172170

R:Sancar, G.B.

Nucleic Acids Res. 13, 8231-8246, 1985

A:Title: Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of the PHR1 phot

A:Reference number: A24046; MUID:86067229; PMID:3906569

A:Accession: A24046

A:Molecule type: DNA

A:Residues: 1-565 <SAN>

A:Cross-references: EMBL:X03183; NID:G4175; PIDN:CAA26944.1; PID:G4176

C:Genetics:

A:Gene: SGB:PHR1

A:Cross-references: SGD:S0005913; MIPS:YOR386w

A:Map position: 15R

C:Superfamily: deoxyribodipyrimidine photo-lyase

C:Keywords: carbon-carbon lyase

Query Match 86.7%; Score 39; DB 2; Length 565;

Best Local Similarity 86.7%; Pred. No. 4.6;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMV 9

Db 86 RLVDNVGLY 94

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 2

S37217

hexon protein - human adenovirus 31 (fragment)

C:Species: Mastadenovirus h31 (human adenovirus 31)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999

C;Accession: S37217
 R;Pring-Akerblom, P.
 submitted to the EMBL Data Library, September 1993
 A;Reference number: S37213
 A;Accession: S37217
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-468 <PRI>
 A;Cross-references: EMBL:X74661; NID:g402765; PIDN:CAA52725.1; PID:g402766
 C;Superfamily: adenovirus hexon protein

Query Match 82.2%; Score 37; DB 2; Length 468;
 Best Local Similarity 66.7%; Pred. No. 9; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMVY 9
 :|||:|
 Db 341 FLYSNVGLY 349

RESULT 3
 VGVNSG
 spike glycoprotein G precursor - sigma virus
 C;Species: sigma virus
 A;Note: host Drosophila melanogaster
 C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
 C;Accession: A27150
 R;Teninges, D.; Bras-Herrens, F.
 J. Gen. Virol. 68, 2625-2638, 1987
 A;Title: Rhabdovirus sigma, the hereditary CO-2 sensitivity agent of Drosophila: nucleocapsid protein
 A;Reference number: A27150; MUID:8803494; PMID:2822842
 A;Accession: A27150
 A;Molecule type: genomic RNA
 A;Residues: 1-526 <TEN>
 A;Cross-references: GB:X06171; NID:g61818; PIDN:CAA29536.1; PID:g61819
 C;Genetics:
 A;Gene: G
 A;Cross-references: FlyBase:FBgn0015809
 C;Superfamily: rhabdovirus spike glycoprotein G
 C;Keywords: glycoprotein; spike protein; transmembrane protein
 F;1-17/Domain: signal sequence #status predicted <Sig>
 F;18-526/Product: spike glycoprotein G #status predicted <SGG>
 F;499-515/Domain: transmembrane #status predicted <TMN>
 F;32,445,459/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.2%; Score 37; DB 1; Length 526;
 Best Local Similarity 66.7%; Pred. No. 11; 0; Indels 0; Gaps 0;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMVY 9
 :|||:|
 Db 350 VLYQSVGMVY 358

RESULT 4
 S33942
 hexon protein - human adenovirus 12
 N;Alternate names: late protein 2
 C;Species: Mastadenovirus h12 (human adenovirus 12)
 C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
 C;Accession: S33942
 R;Springle, J.
 submitted to the EMBL Data Library, June 1993
 A;Reference number: S33928
 A;Accession: S33942
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-919 <SPS>
 A;Cross-references: EMBL:X73487; NID:g313361; PIDN:CAA51891.1; PID:g313376
 C;Superfamily: adenovirus hexon protein

Query Match 82.2%; Score 37; DB 2; Length 919;
 Best Local Similarity 66.7%; Pred. No. 21; 1; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 XLYENVGMVY 9
 :|||:|
 Db 439 FLYSNVGLY 447

RESULT 5
 PL0161
 hemagglutinin - Influenza H2N2 (fragment)
 C;Species: Influenza H2N2
 C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-May-1997
 C;Accession: PL0161
 R;Sweetser, M.T.; Braciale, V.L.; Braciale, T.J.
 J. Exp. Med. 170, 1357-1368, 1989
 A;Title: Class I major histocompatibility complex-restricted T lymphocyte recognition of influenza A virus hemagglutinin
 A;Reference number: PL0161; MUID:90010790; PMID:2477491
 A;Accession: PL0161
 A;Molecule type: mRNA
 A;Residues: 1-20 <SWE>
 A;Experimental source: strain A/JAP/305/57
 C;Comment: this protein plays a major role in initiation of infection and in the pathogenesis of influenza A virus hemagglutinin
 C;Superfamily: influenza virus hemagglutinin
 C;Keywords: hemagglutinin
 F;1-20/Region: immunodominant site recognized by T-lymphocytes

Query Match 80.0%; Score 36; DB 2; Length 20;
 Best Local Similarity 66.7%; Pred. No. 0.45; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMVY 9
 :|||:|
 Db 2 TLYQNVGMVY 10

RESULT 6
 A39365
 cyanamide hydratase (EC 4.2.1.69) - fungus (Myrothecium verrucaria)
 C;Species: Myrothecium verrucaria
 C;Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 15-Sep-2000
 C;Accession: A39365
 R;Maier-Greiner, U.H.; Obermaier-Skrobransky, B.M.M.; Estermaier, L.M.; Kammerloher, W.; F.
 Proc. Natl. Acad. Sci. U.S.A. 88, 4260-4264, 1991
 A;Title: Isolation and properties of a nitrile hydratase from the soil fungus Myrothecium
 A;Reference number: A39365; MUID:91239547; PMID:2034671
 A;Accession: A39365
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-244 <MAI>
 A;Cross-references: GB:M59078; NID:g168392; PIDN:AAA33429.1; PID:g168393
 C;Superfamily: Saccharomyces cerevisiae hypothetical protein YFL061w
 C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 80.0%; Score 36; DB 2; Length 244;
 Best Local Similarity 66.7%; Pred. No. 7.6; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMVY 9
 :|||:|
 Db 169 TLYQNVGMVY 177

RESULT 7
 S39296
 hexon protein - human adenovirus 4
 C;Species: Mastadenovirus h4 (human adenovirus 4)
 C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
 C;Accession: S39296
 R;Pring-Akerblom, P.; Adrian, T.
 submitted to the EMBL Data Library, November 1993
 A;Reference number: S39296
 A;Accession: S39296

A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-447 <PRI>
A:Cross-references: EMBL:X76550; NID:G434903; PIDN:CAA54052.1; PID:G434904
C:Superfamily: adenovirus hexon protein

Query Match 80.0%; Score 36; DB 2; Length 447;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
:|:|:|:|:
Db 355 FLYANVGLY 363

RESULT 8
F70190
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 04-Aug-2003
C:Accession: F70190
R:Prasert, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, son, D.; Peterson, J.; Karlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, son, D.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70190
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-448 <KLE>
A:Cross-references: GB:AE001172; GB:AE000783; NID:G2686654; PIDN:AAC67070.1; PID:G268665
A:Experimental source: strain B31
C:Superfamily: pyrophosphate-dependent phosphofructokinase, Eh/Ppi-PPK type; 6-phosphofructokinase
C:Keywords: Phosphotransferase
F:82-398/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 80.0%; Score 36; DB 1; Length 448;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
:|:|:|:|:
Db 336 LUYEDIGLY 344

RESULT 9
HM1V2
hemagglutinin precursor - influenza A virus (strain A/Japan/305/57[H2])
C:Species: Influenza A virus
A:Variety: strain A/Japan/305/57[H2]
C:Date: 28-Feb-1981 #sequence_revision 28-Feb-1981 #text_change 16-Jul-1999
C:Accession: A04062; S12270
R:Gething, M.J.; Bye, J.; Skehel, J.; Waterfield, M.
Nature 287, 301-306, 1980
A:Title: Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from A/Reference number: A93233; MUID:91030832; PMID:7421990
A:Accession: A04062
A:Molecule type: mRNA
A:Residues: 1-562 <GET>
A:Cross-references: GB:J02127; NID:G324145; PIDN:AAA43185.1; PID:G324146
A:Experimental source: strain A/Japan/305/57[H2]
R:Naeve, C.W.; Williams, D.
EMBO J. 9, 3857-3866, 1990
A:Title: Fatty acids on the A/Japan/305/57 influenza virus hemagglutinin have a role in A:Reference number: S12270; MUID:91065313; PMID:2249653
A:Accession: S12270
A:Molecule type: mRNA
A:Residues: 510-562 <NAE>
A:Experimental source: strain A/Japan/305/57 (H2N2)
C:Superfamily: Influenza virus hemagglutinin
C:Keywords: hemagglutinin; homotrimer; lipoprotein; thiolester bond

F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-339/Product: hemagglutinin chain HA1 #status predicted <HA1>
F:341-562/Product: hemagglutinin chain HA2 #status predicted <HA2>
F:551-558,561/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 80.0%; Score 36; DB 1; Length 562;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
:|:|:|:|:
Db 203 TLYQNVGTY 211

RESULT 10
S57637
hexon protein - human adenovirus 4
C:Species: Mastadenovirus h4 (human adenovirus 4)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C:Accession: S57637
R:Pring-Akerblom, P.; Trigsenaar, J.; Adrian, T.
submitted to the EMBL data Library, February 1995
A:Reference number: S57637
A:Accession: S57637
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-936 <PRI>
A:Cross-references: EMBL:X84646; NID:G986486; PIDN:CAA59139.1; PID:G986487
C:Superfamily: adenovirus hexon protein

Query Match 80.0%; Score 36; DB 2; Length 936;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
:|:|:|:|:
Db 456 FLYANVGLY 464

RESULT 11
B81136
hypothetical protein NMB0968 [imported] - Neisseria meningitidis (strain MC58 serogroup E
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: B81136
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81136
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-29 <RET>
A:Cross-references: GB:AB002448; GB:AB002098; NID:G7226204; PIDN:AAF41373.1; PID:G7226206
C:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0968

Query Match 77.8%; Score 35; DB 2; Length 29;
Best Local Similarity 55.6%; Pred. No. 1.1;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
:|:|:|:|:
Db 21 FLYKNLGLY 29

RESULT 12
H81883
hypothetical protein NMA1165 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C:Species: Neisseria meningitidis
 C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: H81883
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,
 ; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: H81883
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-34 <PAR>
 A:Cross-references: GB:AL162755; GB:AL157959; NID:97379742; PIDN:CAB84427.1; PID:g737988
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMA1165

Query Match 77.8%; Score 35; DB 2; Length 34;
 Best Local Similarity 55.6%; Pred. No. 1.3;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
 :|||:|:
 Db 26 FLYKNLGLY 34

RESULT 13
 A:Species: actin-filament-associated protein 120k form - chicken (fragment)
 C:Species: Gallus gallus (Chicken)
 C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
 C:Accession: A55883
 R:Flynn, D.C.; Kooy, T.C.; Humphries, C.G.; Guappone, A.C.
 J. Biol. Chem. 270, 3894-3899, 1995
 A>Title: AFAP-120. A variant form of the Src SH2/SH3-binding partner AFAP-110 is detected
 A:Reference number: A55883; MUID:95181352; PMID:7876134
 A:Accession: A55883
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-150 <FLY>
 A:Cross-references: GB:L20302

Query Match 77.8%; Score 35; DB 2; Length 150;
 Best Local Similarity 55.6%; Pred. No. 7.1;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
 :|||:|:
 Db 51 MLYDNAGLY 59

RESULT 14
 A:Species: ABC transporter atp-binding protein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
 C:Species: Mycoplasma pulmonis
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
 C:Accession: A99574
 R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
 Nucleic Acids Res. 29, 2145-2153, 2001
 A>Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
 A:Reference number: A99512; MUID:21267165; PMID:11355084
 A:Accession: A99574
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-511 <KUR>
 A:Cross-references: GB:AL445566; PID:g14089911; PIDN:CAC13670.1; GSPDB:GN00153
 A:Experimental source: strain UAB CTIP
 C:Genetics:
 A:Gene: MYPU_4970
 A:Genetic code: SGC3

Query Match 75.6%; Score 34; DB 2; Length 511;
 Best Local Similarity 55.6%; Pred. No. 46;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 XLYENVGMY 9
 :|||:|:
 Db 89 SLYENISVY 97

RESULT 15
 A:Species: myosin I myoA - Emericella nidulans
 C:Species: Emericella nidulans, Aspergillus nidulans
 C>Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 02-Feb-2001
 C:Accession: A56511
 R:McGouldrick, C.A.; Gruver, C.; May, G.S.
 J. Cell Biol. 128, 577-587, 1995
 A>Title: myoA of Aspergillus nidulans encodes an essential myosin I required for secretio
 A:Reference number: A56511; MUID:95164560; PMID:7860631
 A:Accession: A56511
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1249 <MCG>
 A:Cross-references: GB:U12427; NID:9525321; PIDN:AAA67877.1; PID:G525322
 C:Genetics:
 A:Gene: myoA
 C:Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homoclog
 C:Keywords: nucleotide binding; P-loop
 F:53-716/Domain: myosin motor domain homology <VMOT>
 F:143-150/Region: nucleotide-binding motif A (P-loop)
 F:1081-1130/Domain: SH3 homology <SH3>

Query Match 75.6%; Score 34; DB 2; Length 1249;
 Best Local Similarity 55.6%; Pred. No. 1.3e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
 :|||:|:
 Db 1013 DLYQSVGLY 1021

Search completed: July 15, 2004, 07:29:22
 Job time : 13.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:20:47 ; Search time 8 Seconds
(without alignments)
58,579 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLXENVGMY 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|--------------------|
| 1 | 39 | 86.7 | 565 | 1 PHR YEAST | P05066 saccharomyc |
| 2 | 37 | 82.2 | 468 | 1 HEX ADB31 | P36855 human adeno |
| 3 | 37 | 82.2 | 526 | 1 VGLG SIGMA | P18647 sigma virus |
| 4 | 37 | 82.2 | 919 | 1 HEX ADB12 | P19900 human adeno |
| 5 | 36 | 80.0 | 244 | 1 CYAH MYRVE | P22143 myrothecium |
| 6 | 36 | 80.0 | 447 | 1 HEX ADE04 | P36850 human adeno |
| 7 | 36 | 80.0 | 562 | 1 HEMA IAJAP | P03451 influenza a |
| 8 | 35 | 77.8 | 1101 | 1 DIA2 HUMAN | O60879 homo sapien |
| 9 | 33 | 73.3 | 99 | 1 YLM3 CAEEL | P34406 caenorhabdi |
| 10 | 33 | 73.3 | 306 | 1 PYR3 METJA | O58976 methanococc |
| 11 | 33 | 73.3 | 312 | 1 FDH3 HAEIN | P44450 haemophilus |
| 12 | 33 | 73.3 | 313 | 1 CEO2 LACLA | P15244 lactococcus |
| 13 | 33 | 73.3 | 512 | 1 VENV THOGV | P28977 thogoto vir |
| 14 | 33 | 73.3 | 693 | 1 AGLU SULSO | O59645 sulfolobus |
| 15 | 33 | 73.3 | 754 | 1 RAD4 YEAST | P14736 saccharomyc |
| 16 | 32 | 71.1 | 221 | 1 Y805 METJA | O58215 methanococc |
| 17 | 32 | 71.1 | 307 | 1 METF STRLI | O54235 streptomyc |
| 18 | 32 | 71.1 | 437 | 1 PAAK ECOLI | P76085 escherichia |
| 19 | 32 | 71.1 | 450 | 1 DCOR CHICK | P27118 gallus gall |
| 20 | 32 | 71.1 | 455 | 1 DCOR CRIGR | P14019 cricetulus |
| 21 | 32 | 71.1 | 456 | 1 DC02 XENLA | O91884 xenopus lae |
| 22 | 32 | 71.1 | 460 | 1 DCOR XENLA | P27120 xenopus lae |
| 23 | 32 | 71.1 | 461 | 1 DCOR BOVIN | P27117 bos taurus |
| 24 | 32 | 71.1 | 461 | 1 DCOR HUMAN | P11926 homo sapien |
| 25 | 32 | 71.1 | 461 | 1 DCOR MOUSE | P00860 mus musculu |
| 26 | 32 | 71.1 | 461 | 1 DCOR MUSPA | P27119 mus pahari |
| 27 | 32 | 71.1 | 461 | 1 DCOR RAT | P09057 rattus norv |
| 28 | 32 | 71.1 | 519 | 1 ALGG PSEPK | O88nc9 pseudomonas |
| 29 | 32 | 71.1 | 536 | 1 ALGG PSESM | O887q3 pseudomonas |
| 30 | 32 | 71.1 | 593 | 1 PTNB CHICK | O90687 gallus gall |
| 31 | 32 | 71.1 | 593 | 1 PTNB HUMAN | O06124 homo sapien |
| 32 | 32 | 71.1 | 593 | 1 PTNB RAT | P41499 rattus norv |
| 33 | 32 | 71.1 | 671 | 1 RIKI_HUMAN | Q13546 homo sapien |

P52891 saccharomyc
Q12860 homo sapien
P12960 mus musculu
Q63198 rattus norv
P16340 d trifuncti
P37297 saccharomyc
Q09246 caenorhabdi
P45900 bacillus su
P59516 buchnera ap
P75519 mycoplasma
O08333 streptomyc
P43156 hemerocalli

ALIGNMENTS

RESULT 1

PHR_YEAST 726 1 NU84 YEAST
ID PHR_YEAST STANDARD; PRT; 565 AA.
AC P05066;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deoxyribodipyrimidine photolyase, mitochondrial precursor
DE (EC 4.1.99.3) (DNA photolyase) (Photoreactivating enzyme).
GN PHR1 OR YOR386W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86067229; Pubmed=3906569;
RA Sancar G.B.;
RT "Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of
RT the PHR1 photolyase to E. coli photolyase.";
RL Nucleic Acids Res. 13:8231-8246(1985).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=86083177; Pubmed=300886;
RA Yasui A., Langeveld S.A.;
RT "Homology between the photoreactivation genes of Saccharomycetes
RT cerevisiae and Escherichia coli.";
RL Gene 36:349-355(1985).
[3]
RN SEQUENCE FROM N.A.
RA Delius H., Hebling U., Hofmann B.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[4]
RN REVIEW.
RA Sancar G.B., Sancar A.;
RT "Structure and function of DNA photolyases.";
RL Trends Biochem. Sci. 12:259-261(1987).
CC -!- FUNCTION: This enzyme catalyzes the light-dependent monomerization
CC (300-600 nm) of cyclobutyl pyrimidine dimers (in cis-syn
CC configuration), which are formed between adjacent bases on the
CC same DNA strand, upon exposure to ultraviolet radiation.
CC -!- CATALYTIC ACTIVITY: Cyclobutadipyrimidine (in DNA) = 2 pyrimidine
CC residues (in DNA).
CC -!- COFACTOR: Contains 2 chromophores: a reduced flavin (FADH2) and a
CC 5,10-methenyltetrahydrofolate. Both chromophores are bound by non-
CC covalent interactions.
CC -!- SUBCELLULAR LOCATION: Nuclear and mitochondrial.
CC -!- MISCELLANEOUS: This protein belongs to the "short wavelength-type
CC photolyases" with an absorption maximum at about 380 nm.
CC -!- MISCELLANEOUS: There are only 150-300 molecules of photolyase per
CC yeast cell.
CC -!- SIMILARITY: Belongs to the DNA photolyase class-1 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03183; CAA26944.1; -.
DR EMBL; M11578; AAA34875.1; -.
DR EMBL; Z75294; CAA99718.1; -.
DR PIR; S67298; S67298.
DR HSSP; P00314; IJNP.
DR GeronLine; I43974; -.
DR SGD; S0005913; PHR1.
DR InterPro; IPR002081; DNA_photolyase_1.
DR InterPro; IPR006050; DNA_photolyase_N.
DR InterPro; IPR005101; FAD_binding_7.
DR InterPro; IPR006051; FAD_binding_N.
DR Pfam; PF00875; DNA_photolyase; 1.
DR Pfam; PF03441; FAD_binding_7; 1.
DR PRINTS; PR00147; DNAPHOTLYASE.
DR PRODom; PD004390; FAD_binding_N; 1.
DR PROSITE; PS00394; DNA_PHOTOLYASES_1_1; 1.
DR PROSITE; PS00691; DNA_PHOTOLYASES_1_2; 1.
DR Lysase; Chromophore; Flavoprotein_FAD; DNA repair; DNA-binding;
KW Nuclear protein; Mitochondrion; Transic peptide.
FT TRANSIT 1 ? MITOCHONDRION.
FT CHAIN 1 ?
FT DNA_BIND 421 440 H-T-H MOTIF (POTENTIAL).
FT CONFLICT 77 77 V -> A (IN REF. 2).
FT CONFLICT 165 165 T -> S (IN REF. 2).
FT CONFLICT 169 169 S -> T (IN REF. 2).
FT CONFLICT 200 200 D -> S (IN REF. 2).
FT CONFLICT 351 351 S -> R (IN REF. 2).
FT CONFLICT 365 365 G -> E (IN REF. 2).
FT CONFLICT 473 473 E -> K (IN REF. 2).
SQ SEQUENCE 565 AA; 66274 MW; CD4FC3DA6128B97C CRC64;

Query Match 86.78; Score 39; DB 1; Length 565;
Best Local Similarity 66.7%; Pred. No. 3;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLXENVGMY 9
Db :|||:|
86 RLYDNGVLY 94

RESULT 2
HEX_ADE31
ID -HEX_ADE31 STANDARD; PRT; 468 AA.
AC P36855;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hexon protein (late protein 2) (fragment).
GN PII.
OS Human adenovirus type 31.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10529;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VRL 15/62;
RX MEDLINE=94294642; PubMed=8023012;
RA "Type-Akerblom P., Adrian T.;
RT "Type- and group-specific polymerase chain reaction for adenovirus
RT detection."
RL Res. Virol. 145:25-35(1994).
CC -!- FUNCTION: this protein is one of the structural proteins in the
CC viral coat and is synthesized during late infection.
CC -!- SUBUNIT: Homotrimer (by similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74661; CAA52725.1; -.
DR PIR; S37217; S37217.
DR HSSP; P03277; IDEX.
DR InterPro; IPR00736; Adeno_hexon.
DR Pfam; PF01065; Adeno_hexon; 1.
DR PRODom; PD002815; Adeno_hexon; 1.
KW Coat protein; Hexon protein; Late protein.
FT NON_TER 1 468
FT NON_TER 1 468
SQ SEQUENCE 468 AA; 52100 MW; 8727BFA49179CE68 CRC64;

Query Match 82.2%; Score 37; DB 1; Length 468;
Best Local Similarity 66.7%; Pred. No. 6.3;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLXENVGMY 9
Db :|||:|
341 RLYDNGVLY 349

RESULT 3
VGLG_SIGMA
ID -VGLG_SIGMA STANDARD; PRT; 526 AA.
AC P12647;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Spike glycoprotein precursor.
GN G.
OS Sigma virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; unclassified Rhabdoviridae.
OX NCBI_TaxID=11301;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88034947; PubMed=2822842;
RA Teninges D., Bras-Herrang F.;
RT "Rhabdovirus sigma, the hereditary CO2 sensitivity agent of
RT Drosophila: nucleotide sequence of a cDNA clone encoding the
RT Glycoprotein."
RL J. Gen. Virol. 68:2625-2638(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06171; CAA29536.1; -.
DR PIR; A27150; VGVNSG.
DR FlyBase; FBgn0015809; Sigma-Virus\G.
DR InterPro; IPR001903; Rhabd Glycop.
DR Pfam; PF00974; Rhabdo_glycop; 1.
KW Transmembrane; Envelope protein; Glycoprotein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 526 SPIKE GLYCOPROTEIN.
FT CARBOHYD 32 32 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 526 AA; 335607C69249DD9D CRC64;

Query Match 82.2%; Score 37; DB 1; Length 526;
Best Local Similarity 66.7%; Pred. No. 7.1;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLXENVGMY 9
Db :|||:|

```

Db 350 VLYQVGMV 358

RESULT 4

HEX_ADE12 STANDARD; PRT; 919 AA.

AC P19900;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Hexon protein (Late protein 2).

GN PII.

OS Human adenovirus type 12.

OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.

OX NCBI_TaxID=28282;

RN [1]

RP MEDLINE=94076430; PubMed=8254750;

RX STRAIN=Pereira 1131;

RA "Nucleotide sequence of human adenovirus type 12 DNA: comparative functional analysis";

RT J. Virol. 68:379-389 (1994).

RL [2]

RN SEQUENCE OF 888-919 FROM N.A.

RC STRAIN=Pereira 1131;

RX MEDLINE=88303354; PubMed=3043380;

RA Weber J.M., Houde A.;

RT "The primary structure of human adenovirus type 12 protease.";

RL Nucleic Acids Res. 16:7195-7195 (1988).

CC -!- FUNCTION: This protein is one of the structural proteins in the viral coat and is synthesized during late infection.

CC -!- SUBUNIT: Homotrimer (By similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL; X73487; CAA51891.1; -.

DR EMBL; X07655; CAA30501.1; -.

DR EMBL; X07655; CAB37192.1; -.

DR PIR; S01730; S01730.

DR PIR; S33942; S33942.

DR HSP; P03277; 1DHX.

DR InterPro: IPR000736; Adeno_hexon.

DR Pfam; PF01065; Adeno_hexon_1.

DR Pfam; PF03678; Adeno_hexon_C_1.

DR ProDom; PD002815; Adeno_hexon; 1.

KW Coat protein; Hexon protein; Late protein.

SQ SEQUENCE 919 AA; 103039 MW; B37167885A516288 CRC64;

Query Match 82.2%; Score 37; DB 1; Length 919;

Best Local Similarity 66.7%; Pred. No. 13;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVMV 9

Db 439 FLYSNVGLY 447

RESULT 5

CYAH_MYRVE STANDARD; PRT; 244 AA.

AC P22143;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Cyanamide hydratase (EC 4.2.1.69) (Urea hydro-lyase).

GN CAH.

OS Myrothecium verrucaria.

OC Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Hypocreomycetidae; Hypocreales; Mitosporic Hypocreales; Myrothecium.

OX NCBI_TaxID=5532;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=DSM 2087;

RX MEDLINE=91239547; PubMed=2034671;

RA Maier-Greiner U.M., Obermaier-Skrobranek B.M.M., Estermaier L.M., Kammerlocher W., Freund C., Wuefing C., Burkert U.I., Matern D.H., Bauer M., Eulitz M., Kuefrevioglu O.I., Hartmann G.R.;

RT "Isolation and properties of a nitrile hydratase from the soil fungus Myrothecium verrucaria that is highly specific for the fertilizer cyanamide and cloning of its gene";

RL Proc. Natl. Acad. Sci. U.S.A. 88:4260-4264 (1991).

CC -!- FUNCTION: When used as herbicide in agriculture, cyanamide can be transformed, after sowing, in soil fertilizing ammonia by the combined action of M.verrucaria cyanamide hydratase and urease.

CC -!- CATALYTIC ACTIVITY: Urea = cyanamide + H(2)O.

CC -!- COFACTOR: Zinc.

CC -!- SUBUNIT: Homohexamer.

CC -!- MISCELLANEOUS: This enzyme is highly specific for cyanamide.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL; M59078; AAA33429.1; -.

DR PIR; A39365; A39365.

DR InterPro: IPR006674; HD.

DR InterPro: IPR003607; Met_phosphohydro.

DR Pfam; PF01966; HD; 1.

DR SMART; SM00471; HDC; 1.

KW Lyase; Zinc.

SQ SEQUENCE 244 AA; 26966 MW; 880FALL1F30E31CE2 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 244;

Best Local Similarity 66.7%; Pred. No. 5;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVMV 9

Db 169 FLYDNVGV 177

RESULT 6

HEX_ADE04 STANDARD; PRT; 447 AA.

AC P36850;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Hexon protein (Late protein 2) (Fragment).

GN PII.

OS Human adenovirus type 4.

OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.

OX NCBI_TaxID=28280;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Isolate RJ-67;

RX MEDLINE=95407102; PubMed=7676636;

RA Pring-Akerblom P., Trijssenaar J., Adrian T.;

RT "Sequence characterization and comparison of human adenovirus subgenus B and E hexons";

RL Virology 212:232-236 (1995).

CC -!- FUNCTION: This protein is one of the structural proteins in the viral coat and is synthesized during late infection.

CC -!- SUBUNIT: Homotrimer (By similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X76550; CAA54052.1; -
 DR PIR; S39296; S39296.
 DR HSSP; P03277; LDHX.
 DR InterPro; IPR000736; Adeno_hexon.
 DR Pfam; PF01065; Adeno_hexon.1.
 DR ProDom; PD002815; Adeno_hexon; 1.
 KW Coat protein; Hexon protein; Late protein.
 FT NON_TER 1 447
 FT TER 447 447
 SQ SEQUENCE 447 AA; 49553 MW; A7AE1977F707BD4D CRC64;

Query Match 80.0%; Score 36; DB 1; Length 447;
 Best Local Similarity 66.7%; Pred. No. 9.6;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMV 9
 :||:||||
 Db 355 FLYANVGLY 363

RESULT 7

ID HEMA IAJAP STANDARD; PRT; 562 AA.
 AC P03451;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hemagglutinin precursor [contains: Hemagglutinin HAI chain;
 DE Hemagglutinin H2A chain].
 GN HA.
 OS Influenza A virus (strain A/Japan/305/57).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Influenza A viruses; Influenzavirus A.
 OX NCBI_TaxID=11421;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=81030852; PubMed=7421990;
 RA Gething M.-J., Bye J., Skehel J.J., Waterfield M.;
 RA "Cloning and DNA sequence of double-stranded copies of haemagglutinin
 RA genes from H2 and H3 strains elucidates antigenic shift and drift in
 RA human influenza virus".
 RL Nature 287:301-306(1980).

CC -!- FUNCTION: Hemagglutinin is responsible for attaching the virus to
 CC cell receptors and for initiating infection.

CC -!- SUBUNIT: Homotrimer. Each of the monomers is formed by two chains
 CC (HA1 and HA2) linked by a disulfide bond.

CC -!- SIMILARITY: Belongs to the influenza viruses hemagglutinin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J02127; AAA43185.1; -

DR PIR; A04062; HMI2.

DR HSSP; P03437; LHTM.

DR InterPro; IPR008980; Capsid_hemag.

DR InterPro; IPR001364; Hemagglutn.

DR Pfam; PF00509; Hemagglutinin; 1.

DR PRINTS; PR00329; HEMAGGLUTIN12.

DR ProDom; PD000225; Hemagglutn; 1.

DR Envelope protein; Hemagglutinin; Glycoprotein; Signal.

FT SIGNAL 1 15

FT CHAIN 16 339 HEMAGGLUTININ HAI CHAIN.
 FT CHAIN 341 562 HEMAGGLUTININ HA2 CHAIN.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 562 AA; 63118 MW; 6B7FDC0C38993630 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 562;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMV 9
 :||:||||
 Db 203 TLYQNVGT 211

RESULT 8

ID DIA2 HUMAN STANDARD; PRT; 1101 AA.
 AC O60879; O60879; Q9UUL2;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Diaphanous protein homolog 2 (Diaphanous-related formin 2) (DRP2).

GN DIAHP2 OR DIA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RX MEDLINE=98163437; PubMed=9497258;

RA Bione S., Sala C., Manzini C., Arrigo G., Zuffardi O., Banfi S.,

RA Borsani G., Jonveaux P., Philippe C., Zuccotti M., Ballabio A.,

RA Toniolo D.;

RT "A human homologue of the Drosophila melanogaster diaphanous gene is
 RT disrupted in a patient with premature ovarian failure: evidence for
 RT conserved function in oogenesis and implications for human
 RT sterility".

RL Am. J. Hum. Genet. 62:533-541(1998).

RN [2]

RP SEQUENCE OF 685-906 FROM N.A.

RA Heath P.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: May be involved in oogenesis.

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=2;

CC Name=DIA-156;

CC IsoId=O60879-1; Sequence=Displayed;

CC Name=DIA-12C;

CC IsoId=O60879-2; Sequence=VSP_001573;

CC -!- TISSUE SPECIFICITY: Expressed in testis, ovary, small intestine,
 CC prostate, lung, liver, kidney, Leukocytes

CC -!- DEVELOPMENTAL STAGE: Expressed from E16 in ovary and testis and
 CC during P6-P16 during differentiation of ovarian follicles.

CC -!- DOMAIN: DRPs are regulated by intramolecular GBD-DAD binding where
 CC Rho-GTP activates the DRPs by disrupting the GBD-DAD interaction
 CC (By similarity).

CC -!- DISEASE: Defects in DIAHP2 are a cause of premature ovarian
 CC failure (POF) [MIM:311360].

CC -!- SIMILARITY: Contains 1 GTPase-binding (GBD) domain.

CC -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.

CC -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.

CC -!- SIMILARITY: Contains 1 Formin homology 3 (FH3) domain.

CC -!- SIMILARITY: Contains 1 DRF autoregulatory (DAD) domain.

CC -!- SIMILARITY: Belongs to the formin homology family. Diaphanous
 CC subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y15909; CAA75870.1; --
 EMBL; Y15908; CAA75869.1; --
 EMBL; AL031053; CAB39108.1; --
 MIM; 300108; --
 MIM; 311360; --
 GO; GO:0005102; P:receptor binding; TAS.
 GO; GO:0016288; P:cytokinesis; TAS.
 GO; GO:0007292; P:female gamete generation; TAS.
 InterPro; IPR003104; FH2.
 Pfam; PF02181; FH2; 1.
 SMART; SM00498; FH2; 1.
 Alternative splicing; Coiled coil; Repeat.
 FT DOMAIN 86 285
 FT DOMAIN 184 482
 FT DOMAIN 366 418
 FT DOMAIN 487 547
 FT DOMAIN 549 623
 FT DOMAIN 628 1071
 FT DOMAIN 903 1053
 FT DOMAIN 1054 1068
 FT DOMAIN 1072 1075
 FT DOMAIN 257 260
 FT DOMAIN 543 546
 FT DOMAIN 562 572
 FT DOMAIN 576 585
 FT DOMAIN 591 597
 FT DOMAIN 603 608
 FT DOMAIN 613 616
 FT DOMAIN 1038 1041
 FT VARSPLIC 1081 1101
 SEQUENCE 1101 AA; 125568 MW; 399FLC292D79188B CRC64;

Query Match 77.8%; Score 35; DB 1; Length 1101;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
 :|||||:
 Db 970 KLYENLGEY 978

RESULT 9
 ID YLW3 CAEEL STANDARD; PRT; 99 AA.
 AC P34406;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein F22B7.3 in chromosome III.
 GN F22B7.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Andersson K., Baynes C., Berks M., Coulson A.,
 Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., M.,
 Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,

Johnston L., Jones M., Kersey J., Kirsten J., Laister N.,
 Lattelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 Parsons J., Percy C., Rifkin L., Ropra A., Saunders D., Showken R.,
 Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden K.,
 Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
 Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; L12018; AAA5463.1; --
 PIR; S44632; S44632.
 WormPep; F22B7.3; CE00156.
 KW Hypothetical protein.
 SQ SEQUENCE 99 AA; 11665 MW; 78FC94DBD3C8B585 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 99;
 Best Local Similarity 71.4%; Pred. No. 8.1;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YENVGMY 9
 :|||||:
 Db 21 YENLGMF 27

RESULT 10
 ID PYRB METJA STANDARD; PRT; 306 AA.
 AC Q58976;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartate carbamoyltransferase (EC 2.1.3.2) (Aspartate
 transcarbamylase) (ATCase).
 DE PYRB OR MJ1581.
 GN Methanococcus jannaschii.
 OS Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 Overbeek R., Kirkness B.F., Weinstock K.G., Merrick J.M., Glodek A.,
 Scott J.L., Geoghegan N.S.M., Weidman J.F., Sadow P.W., Hanna M.C.,
 Utterback T.R., Kelley J.M., Peterson J.D., Burdick D., Borodovsky M.,
 Corton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=20283607; PubMed=10748119;
 RA Hack E.S., Vorobyova T., Sakash J.B., West J.M., Macol C.P., Herve G.,
 Williams M.K., Kantrowitz E.R.;
 RT "Characterization of the aspartate transcarbamoylase from
 Methanococcus jannaschii.";
 RL J. Biol. Chem. 275:15820-15827(2000).
 RN [3]
 RP CRYSTALLIZATION, AND X-RAY CRYSTALLOGRAPHY.

```

RA MEDLINE=20402716; PubMed=10944354;
RA Vitali J., Vorobyova T., Webster G., Kantrowitz E.R.;
RT "Crystallization and structure determination of the catalytic trimer
RT of Methanococcus jannaschii aspartate transcarbamoylase.";
RL Acta Crystallogr. D 56:1061-1063(2000).
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
CC + N-carbamoyl-L-aspartate.
CC -!- PATHWAY: Pyrimidine biosynthesis; second step.
CC -!- SUBUNIT: HETEROODIMER (2C3:3E2) OF SIX CATALYTIC PYR CHAINS
CC ORGANIZED AS TWO TRIMERS (C3), AND SIX REGULATORY PYR CHAINS
CC ORGANIZED AS THREE DIMERS (R2).
CC -!- SIMILARITY: Belongs to the Afcase/OTCase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U67598; AAB99601.1; --
DR PIR; D64497; D64497.
DR HSSP; P00479; 3CSU.
DR TIGR; MJ1581; --.
DR HAMAP; MF_00001; --; 1.
DR InterPro; IPR006130; Asp/Orn_Cotransf.
DR InterPro; IPR002082; Asp_carbMltransf.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00100; AOTCASE.
DR TIGRFAMs; TIGR00670; asp_carb tr; 1.
DR PROSITE; PS0097; CARBAMOYLTRANSFERASE; 1.
DR Pyrimidine biosynthesis; Transferrase; Complete proteome.
KW SEQUENCE 306 AA; 35159 MW; CBDC31FC450CEP6A CRC64;
-----
Query Match 73.3%; Score 33; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLVENVGMY 9
Db 174 SLFENVEMY 182
:::|::|::|

RESULT 11
FDXH HAEIN
ID _FDXH HAEIN STANDARD; PRT; 312 AA.
AC P44450;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase, iron-sulfur subunit (Formate dehydrogenase beta
DE subunit) (FDH beta subunit).
GN FDH OR HI0007.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
CX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-L., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

```

```

RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: ALLOWS TO USE FORMATE AS MAJOR ELECTRON DONOR DURING
CC ANAEROBIC RESPIRATION. THE BETA CHAIN IS AN ELECTRON TRANSFER UNIT
CC CONTAINING 4 CYSTEINE CLUSTERS INVOLVED IN THE FORMATION OF IRON-
CC SULFUR CENTRES. ELECTRONS ARE TRANSFERRED FROM THE GAMMA CHAIN TO
CC THE MOLYBDENUM COPROFACTOR OF THE ALPHA SUBUNIT (BY SIMILARITY).
CC -!- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
CC BY SUBUNITS ALPHA, BETA AND GAMMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: ORTHOLOG OF BOTH E.COLI FDH AND FDH.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U32686; AAC21685.1; --
DR PIR; A64042; A64042.
DR HSSP; P00193; 1DUR.
DR TIGR; HI0007; --.
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR InterPro; IPR006470; FDH_beta.
DR Pfam; PF00037; fer4; 1.
DR TIGRFAMs; TIGR01582; FDH-beta; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW Electron transport; 4Fe-4S; Iron-sulfur; Transmembrane;
KW Complete proteome.
FT METAL 44 44 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 47 47 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 50 50 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 54 54 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 106 106 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 109 109 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 114 114 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 118 118 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 139 139 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 142 142 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 145 145 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 149 149 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 166 166 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT METAL 169 169 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT METAL 181 181 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT METAL 185 185 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 312 AA; 34068 MW; AA49DD3C17064866 CRC64;
-----
Query Match 73.3%; Score 33; DB 1; Length 312;
Best Local Similarity 71.4%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YENVGMY 9
Db 214 YENAGLY 220
|||::|::|

RESULT 12
CE02 LACLA
ID _CE02 LACLA STANDARD; PRT; 313 AA.
AC P15244;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE N(5)-(carboxyethyl)ornithine synthase (EC 1.5.1.24) (N(5)-(L-1-
DE carboxyethyl)-L-ornithine:NADP(+) oxidoreductase) (CEOS).
GN CEO.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
CX NCBI_TaxID=1360;

```

```

RN  SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-15.
RP  STRAIN=K1-23; TRANSPOS=tn5306;
RC  MEDLINE=95263576; PubMed=7744873;
RA  Donkersloot J.A., Thompson J.;
RT  "Cloning, expression, sequence analysis, and site-directed
RT  mutagenesis of the Tn5306-encoded N5-(carboxyethyl)ornithine synthase
RT  from Lactococcus lactis K1.";
RL  J. Biol. Chem. 270:12226-12234(1995).
RN  [2]
RP  SEQUENCE OF 1-37.
RC  STRAIN=K1;
RX  MEDLINE=89255467; PubMed=2498334;
RA  Thompson J.;
RT  "N5-(L-1-carboxyethyl)-L-ornithine:NADP+ oxidoreductase from
RT  Streptococcus lactis. Purification and partial characterization.";
RL  J. Biol. Chem. 264:9592-9601(1989).
RN  [3]
RP  SEQUENCE OF 256-263, AND CHARACTERIZATION.
RC  STRAIN=K1;
RX  MEDLINE=20014035; PubMed=10548058;
RA  Sackett D.L., Ruvinov S.B., Thompson J.;
RT  "N5-(L-1-carboxyethyl)-L-ornithine synthase: physical and spectral
RT  characterization of the enzyme and its unusual low pKa fluorescent
RT  tyrosine residues.";
RL  Protein Sci. 8:2121-2129(1999).
RN  [4]
RP  FOLDING STUDIES.
RC  STRAIN=K1;
RX  MEDLINE=99456521; PubMed=10525296;
RA  Ruvinov S.B., Thompson J., Sackett D.L., Ginsburg A.;
RT  "Tetrameric N(5)-(L-1-carboxyethyl)-L-ornithine synthase: guanidine.
RT  HCl-induced unfolding and a low temperature requirement for
RT  refolding.";
RL  Arch. Biochem. Biophys. 371:115-123(1999).
CC  -1- CATALYTIC ACTIVITY: N(5)-(L-1-carboxyethyl)-L-ornithine + NADP(+)
CC  + H(2)O = L-ornithine + pyruvate + NADPH.
CC  -1- SUBUNIT: Homotetramer.
CC  -1- MASS SPECTROMETRY: MW=35.355; METHOD=MALDI.
CC  -1- MISCELLANEOUS: In the reverse direction L-lysine can act instead
CC  of L-ornithine, more slowly, yielding N(6)-(L-1-carboxyethyl)-L-
CC  lysine.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: U23376; AAA96385.1; --
DR  FIR: A57499; A57499.
DR  InterPro: IPR007698; Aladh_PNT_C.
DR  InterPro: IPR007886; Aladh_PNT_N.
DR  Pfam: PF01262; Aladh_PNT_C; 1.
DR  Pfam: PF05222; Aladh_PNT_N; 1.
KW  Oxidoreductase; NADP-
FT  NP_BIND 171 176 NADPH (POTENTIAL).
FT  MOTAGEN 15 15 R->K: LOSS OF ACTIVITY.
SQ  SEQUENCE 313 AA; 35323 MW; B17FE0F477113C77 CRC64;
Query Match 73.3%; Score 33; DB 1; Length 313;
Best Local Similarity 55.6%; Pred.No. 27;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 XLYENVGMY 9
DB 261 PIYENAGKY 269
RESULT 13
VENV_THOGV

```

```

ID  VENV_THOGV STANDARD; PRT; 512 AA.
AC  P28977;
DT  01-DEC-1992 (Rel. 24, Created)
DT  01-DEC-1992 (Rel. 24, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Envelope glycoprotein precursor (Surface glycoprotein 75).
GN  PA.
OS  Thogoto virus (isolate SiAr 126) (Tho).
OC  Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC  Thogotovirus.
OX  NCBI_TaxID=126796;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92124738; PubMed=1733105;
RA  Morse M.A., Marriott A.C., Nuttall P.A.;
RT  "The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like
RT  virus) is related to the baculovirus glycoprotein GP64.";
RL  Virology 186:640-646(1992).
CC  -1- FUNCTION: POSSIBLE ROLE IN ENDOCYTOTIC FUSION EVENTS DURING
CC  INFECTION.
CC  -1- SUBUNIT: Monomer (Probable)
CC  -1- SIMILARITY: TO DHORI VIRUS ENVELOPE GLYCOPROTEIN AND TO
CC  BACULOVIRUSES MAJOR ENVELOPE GLYCOPROTEIN (P64/P67).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M77280; AAA47918.1; --
DR  FIR: A40821; VGIIVTH.
DR  InterPro: IPR004955; Baculo_gp64.
DR  Pfam: PF03273; Baculo_gp64; 1.
KW  Glycoprotein; Transmembrane; Signal.
FT  SIGNAL 1 15 POTENTIAL.
FT  CHAIN 16 512 ENVELOPE GLYCOPROTEIN.
FT  TRANSMEM 479 502 POTENTIAL.
FT  CARBOHYD 185 185 N-LINKED (GLCNAC...) (POTENTIAL).
FT  CARBOHYD 263 263 N-LINKED (GLCNAC...) (POTENTIAL).
FT  CARBOHYD 289 289 N-LINKED (GLCNAC...) (POTENTIAL).
FT  CARBOHYD 378 378 N-LINKED (GLCNAC...) (POTENTIAL).
FT  CARBOHYD 416 416 N-LINKED (GLCNAC...) (POTENTIAL).
SQ  SEQUENCE 512 AA; 57550 MW; 0398FC36284A0DF1 CRC64;
Query Match 73.3%; Score 33; DB 1; Length 512;
Best Local Similarity 55.6%; Pred.No. 45;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 XLYENVGMY 9
DB 484 LLYGNIGVY 492
RESULT 14
AGLU_SULSO
ID  AGLU_SULSO STANDARD; PRT; 693 AA.
AC  O59645;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-glucosidase (EC 3.2.1.20) (Maltase).
GN  MALA OR SSO3051 OR C23_036.
OS  Sulfolobus solfataricus.
OC  Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC  Sulfolobus.
OX  NCBI_TaxID=2287;
RN  [1]
RP  SEQUENCE FROM N.A., AND SEQUENCE OF 1-20 AND 552-561.
RC  STRAIN=98/2;
RX  MEDLINE=98155158; PubMed=9495770;

```

RA Rolfsmeier M., Haseltine C., Bini E., Clark A., Blum P.:
 RT "Molecular characterization of the alpha-glucosidase gene (mala) from
 the hyperthermophilic archaeon Sulfolobus solfataricus.";
 RN J. Bacteriol. 180:1287-1295(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri P., Zivanovic Y., Allard G.,
 RA Awayez M.J., Chan-Weiner C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.B., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -!- FUNCTION: MAJOR SOLUBLE ALPHA-GLUCOSIDASE.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
 CC linked D-glucose residues with release of D-glucose.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- INDUCTION: EXPRESSED DURING GROWTH ON MALTOSSE.
 CC -!- MISCELLANEOUS: THE PH OPTIMUM FOR MALTOSSE HYDROLYSIS IS 4.5, AND
 CC 5.5 FOR GLUCOGEN HYDROLYSIS.
 CC -!- SIMILARITY: Belongs to family 31 of glycosyl hydrolases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF042494; AAC38215.1; -;
 CC EMBL; AE006896; AAK43151.1; -;
 CC PIR; H90486; H90486.
 CC InterPro; IPR000322; Glyco_hydro_31.
 CC Pfam; PF01085; Glyco_hydro_31; 1.
 CC PROSITE; PS00129; GLYCOSYL_HYDROL_F31_1; 1.
 CC PROSITE; PS00707; GLYCOSYL_HYDROL_F31_2; FALSE_NEG.
 CC Hydrolyase; Glycosidase; Complete proteome.
 CC ACT SITE 320 320 BY SIMILARITY
 CC SEQUENCE 693 AA; 80441 MW; 2789952COAYB3858 CRC64;
 CC
 CC Query Match 73.3%; Score 33; DB 1; Length 693;
 CC Best Local Similarity 55.6%; Pred. No. 62;
 CC Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 XLYENVGMY 9
 CC Db 5 KIYENKGVY 13
 CC
 CC RESULT 15
 CC RAD4 YEAST
 CC ID RAD4 YEAST STANDARD; PRT; 754 AA.
 CC P14736;
 CC DT 01-APR-1990 (Rel. 14, Created)
 CC DT 01-APR-1990 (Rel. 14, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE DNA repair protein RAD4.
 CC GN RAD4 OR YER162C.
 CC OS Saccharomyces cerevisiae (Baker's yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC NCBI_TaxID=4932;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RX MEDLINE=89232744; PubMed=3073107;
 CC Gietz R.D., Prakash S.;
 RT "Cloning and nucleotide sequence analysis of the Saccharomycetes
 RT cerevisiae RAD4 gene required for excision repair of UV-damaged

RT DNA.";
 RL Gene 74:535-541(1988).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=89197751; PubMed=2649477;
 RA Couto L.B., Friedberg E.C.;
 RT "Nucleotide sequence of the wild-type RAD4 gene of Saccharomycetes
 RT cerevisiae and characterization of mutant rad4 alleles.";
 RL J. Bacteriol. 171:1862-1869(1989).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313264; PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 RA Hunkeler-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomycetes cerevisiae chromosome V.";
 RL Nature 387:78-81(1997).
 CC -!- FUNCTION: Involved in nucleotide excision repair of DNA damaged
 CC with UV light, bulky adducts, or cross-linking agents.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the XPC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M26050; AAA34944.1; -;
 CC EMBL; M24928; AAA34945.1; -;
 CC EMBL; U18917; AAB64689.1; -;
 CC PIR; S30814; DDBYD4.
 CC GERMOnline; I39239; -;
 CC SGD; S0000964; RAD4.
 CC GO; GO:0000111; C:nucleotide excision repair factor 2 complex; IDA.
 CC GO; GO:0000108; C:repairosome; IDA.
 CC GO; GO:0003684; F:damaged DNA binding; IDA.
 CC InterPro; IPR004583; Rad4.
 CC Pfam; PF03835; Rad4; 1.
 CC TIGRFAMs; TIGR00605; rad4; 1.
 CC DNA repair; DNA-binding; Nuclear protein.
 CC DNA BIND 250 269 POTENTIAL.
 CC CONFLICT 223 225 VGI -> EGL (IN REF. 3).
 CC SEQUENCE 754 AA; 87174 MW; 788C146DC4BD2BF8 CRC64;
 CC
 CC Query Match 73.3%; Score 33; DB 1; Length 754;
 CC Best Local Similarity 71.4%; Pred. No. 68;
 CC Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 3 YENVGMY 9
 CC Db 220 YDNVGIY 226
 CC
 CC Search completed: July 15, 2004, 07:27:01
 CC Job time : 10 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:25:27 ; Search time 33 seconds
(without alignments)
86.050 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLXENVGMVY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mic.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|-----------|--------------------|
| 1 | 37 | 82.2 | 540 | 12 Q88452 | Q88452 sigma virus |
| 2 | 37 | 82.2 | 914 | 12 Q9IF30 | Q9IF30 bovine aden |
| 3 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 4 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 5 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 6 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 7 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 8 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 9 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 10 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 11 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 12 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 13 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 14 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 15 | 36 | 80.0 | 339 | 12 Q9IF30 | Q9IF30 influenza a |
| 16 | 36 | 80.0 | 339 | 12 Q997B2 | Q997B2 influenza a |

| | | | | | |
|----|----|------|-----|-----------|--------------------|
| 17 | 36 | 80.0 | 359 | 12 Q997B3 | Q997B3 influenza a |
| 18 | 36 | 80.0 | 359 | 12 Q997B4 | Q997B4 influenza a |
| 19 | 36 | 80.0 | 359 | 12 Q997B1 | Q997B1 influenza a |
| 20 | 36 | 80.0 | 373 | 12 Q9WQX2 | Q9WQX2 influenza a |
| 21 | 36 | 80.0 | 376 | 12 Q9WQX1 | Q9WQX1 influenza a |
| 22 | 36 | 80.0 | 376 | 12 Q9WQW1 | Q9WQW1 influenza a |
| 23 | 36 | 80.0 | 376 | 12 Q9WQW4 | Q9WQW4 influenza a |
| 24 | 36 | 80.0 | 378 | 12 Q9WQX0 | Q9WQX0 influenza a |
| 25 | 36 | 80.0 | 378 | 12 Q9WQX8 | Q9WQX8 influenza a |
| 26 | 36 | 80.0 | 378 | 12 Q9WQW6 | Q9WQW6 influenza a |
| 27 | 36 | 80.0 | 378 | 12 Q9WQW2 | Q9WQW2 influenza a |
| 28 | 36 | 80.0 | 379 | 12 Q9WQX3 | Q9WQX3 influenza a |
| 29 | 36 | 80.0 | 380 | 12 Q9WQV9 | Q9WQV9 influenza a |
| 30 | 36 | 80.0 | 381 | 12 Q9WQW7 | Q9WQW7 influenza a |
| 31 | 36 | 80.0 | 381 | 12 Q9WQW5 | Q9WQW5 influenza a |
| 32 | 36 | 80.0 | 381 | 12 Q9WQW3 | Q9WQW3 influenza a |
| 33 | 36 | 80.0 | 381 | 12 Q9WQW0 | Q9WQW0 influenza a |
| 34 | 36 | 80.0 | 448 | 16 OS1669 | OS1669 borrelia bu |
| 35 | 36 | 80.0 | 560 | 12 Q9WQW9 | Q9WQW9 influenza a |
| 36 | 36 | 80.0 | 562 | 12 Q67032 | Q67032 influenza a |
| 37 | 36 | 80.0 | 562 | 12 Q67085 | Q67085 influenza a |
| 38 | 36 | 80.0 | 562 | 12 Q67208 | Q67208 influenza a |
| 39 | 36 | 80.0 | 562 | 12 Q67120 | Q67120 influenza a |
| 40 | 36 | 80.0 | 562 | 12 Q67101 | Q67101 influenza a |
| 41 | 36 | 80.0 | 562 | 12 Q67284 | Q67284 influenza a |
| 42 | 36 | 80.0 | 562 | 12 Q67165 | Q67165 influenza a |
| 43 | 36 | 80.0 | 562 | 12 Q67143 | Q67143 influenza a |
| 44 | 36 | 80.0 | 562 | 12 Q67140 | Q67140 influenza a |
| 45 | 36 | 80.0 | 562 | 12 Q67326 | Q67326 influenza a |

ALIGNMENTS

RESULT 1

Q88452 PRELIMINARY; PRT; 540 AA.
AC Q88452; DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Glycoprotein.
GN G.
OS Sigma virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; unclassified Rhabdoviridae.
OX NCBI_TaxID=11301;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=234HRC;
RX MEDLINE=96074506; PubMed=7491755;
RA Landes-Devauchelle C., Bras F., Derelee S., Teninges D.;
RT "Gene 2 of the sigma rhabdovirus genome encodes the P protein, and
RT gene 3 encodes a protein related to the reverse transcriptase of
RT retroelements.";
RL Virology 213:300-312(1995).
DR EMBL: X91062; CAA62517.1; .
DR InterPro: IPR001903; Rhabd_glycop.
DR Pfam: PF00974; Rhabdo_Glycop; 1.
SQ SEQUENCE 540 AA; 60771 MW; 7A0B553D1EA5E98A CRC64;

Query Match 82.2%; Score 37; DB 12; Length 540;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLXENVGMVY 9
Db 364 VLYQSVGMVY 372

RESULT 2

Q9IF30 PRELIMINARY; PRT; 914 AA.
ID Q9IF30

```

AC Q9IF30;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hexon protein (Fragment).
OS Bovine adenovirus type 10 (Mastadenovirus bos10).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=39788;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=78-5371;
RA Leimkuhl H.D., Hobbs L.A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282774; AAF82136.1; -.
DR HSP; P03277; IDHX.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000736; Adeno_hexon.
DR Pfam; PF01085; Adeno_hexon; 1.
DR Pfam; PF03678; Adeno_hexon_C; 1.
DR ProDom; PD002815; Adeno_hexon; 1.
FT NON_TER 914 914
SQ SEQUENCE 914 AA; 103905 MW; 5508E006957739CD CRC64;

Query Match 82.2%; Score 37; DB 12; Length 914;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
DB 433 FLYSNUGLY 441

RESULT 3
Q9IFF9 PRELIMINARY; PRT; 339 AA.
AC Q9IFF9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Davis/1/57(H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220951;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=A/Davis/1/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RL "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270719; AAF82103.1; -.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 339 339
SQ SEQUENCE 339 AA; 37810 MW; 7D35925ED7538B08 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
DB 203 FLYSNUGLY 211

RESULT 5
Q9IFF2 PRELIMINARY; PRT; 339 AA.
AC Q9IFF2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Victoria/15681/59(H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220956;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=A/Victoria/15681/59;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RL "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

```

```

QY 1 XLYENVGMY 9
DB 203 TLYQNVGTY 211

RESULT 4
Q9IFF4 PRELIMINARY; PRT; 339 AA.
AC Q9IFF4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Malaya/16/58(H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220954;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=A/Malaya/16/58;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RL "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270724; AAF82108.1; -.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 339 339
SQ SEQUENCE 339 AA; 37893 MW; D59A261E1EB9B621 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
DB 203 TLYQNVGTY 211

RESULT 5
Q9IFF2 PRELIMINARY; PRT; 339 AA.
AC Q9IFF2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Victoria/15681/59(H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220956;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=A/Victoria/15681/59;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RL "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

```

```
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270726; AAF82110.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; Hemagglutn12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37964 MW; 97239D60CD1FFD08 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 6
Q9IFF6 PRELIMINARY; PRT; 339 AA.
AC Q9IFF6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/R1/5+/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135328;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=A/R1/5+/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270722; AAF82106.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; Hemagglutn12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37853 MW; 7C70576EBB5B2FC0 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 7
Q9IFF0 PRELIMINARY; PRT; 339 AA.
AC Q9IFF0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Chile/6/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135323;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=A/Chile/6/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270721; AAF82105.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
```

```
Q9IFF0 PRELIMINARY; PRT; 339 AA.
AC Q9IFF0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Chile/6/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135323;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=A/Chile/6/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270728; AAF82112.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; Hemagglutn12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37810 MW; 7D35925ED7538B08 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 8
Q9IFF7 PRELIMINARY; PRT; 339 AA.
AC Q9IFF7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (strain A/Ann Arbor/6/60).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135322;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=A/Ann Arbor/6/60;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270721; AAF82105.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
```

```

DR InterPro: IPR001364; Hemagglutn.
DR Pfam: PF00509; Hemagglutinin. 1.
DR PRINTS: PR00329; HEMAGGLUTN12.
DR ProDom: PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37896 MW; FECE7718D2628F0E CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 9
Q91FF8 PRELIMINARY; PRT; 339 AA.
AC Q91FF8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Albany/7/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135321;
RN [1];
RP SEQUENCE FROM N.A.
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RA "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL: AF270720; AAF82104.1; -.
DR GO: GO:0019031; C:viral envelope; IEA.
DR InterPro: IPR008980; Capsid_hemag.
DR Pfam: PF00509; Hemagglutinin; 1.
DR PRINTS: PR00329; HEMAGGLUTN12.
DR ProDom: PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37753 MW; 2ADC4BA8C590ADCE CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 11
Q91FF8 PRELIMINARY; PRT; 339 AA.
AC Q91FF8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/RI/5-/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135329;
RN [1];
RP SEQUENCE FROM N.A.
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RA "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL: AF270718; AAF82102.1; -.
DR GO: GO:0019031; C:viral envelope; IEA.
DR InterPro: IPR008980; Capsid_hemag.
DR Pfam: PF00509; Hemagglutinin; 1.
DR PRINTS: PR00329; HEMAGGLUTN12.
DR ProDom: PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37798 MW; FE7698C4DC1D15E6 CRC64;

```

```
Query Match      80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 12
Q9IFP3 PRELIMINARY; PRT; 339 AA.
AC Q9IFP3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Sao Paulo/3/59 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135330;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Sao Paulo/3/59;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270725; AAF82109.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37895 MW; 97D69D60CD5AFD08 CRC64;

Query Match      80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 13
Q9IFP1 PRELIMINARY; PRT; 339 AA.
AC Q9IFP1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Ohio/2/59 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135327;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Ohio/2/59;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
```

```
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270727; AAF82111.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37991 MW; F6BC8A0403FD40CC CRC64;

Query Match      80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 203 TLYQNVGT 211

RESULT 14
Q9IFG2 PRELIMINARY; PRT; 339 AA.
AC Q9IFG2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/El Salvador/2/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135325;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/El Salvador/2/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270716; AAF82100.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37874 MW; 237050D99292320A CRC64;

Query Match      80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMY 9
Db 203 TLYQNVGT 211
```

RESULT 15

```

Q9IFG1
ID Q9IFG1 PRELIMINARY; PRT; 339 AA.
AC Q9IFG1
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hemagglutinin (Fragment)
OS Influenza A virus (A/Leningrad/134/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=128148;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Leningrad/134/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270717; AAF82101.1; -
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 339
SQ SEQUENCE 339 AA; 37825 MW; EC97187675C23218 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 339;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
Db 203 TLXQNVGTY 211

```

Search completed: July 15, 2004, 07:30:40
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:21:32 ; Search time 48 Seconds

(without alignments)
52.978 Million cell updates/sec

Title: SEQIMOD

Perfect score: 39

Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No | Score | Query Match | Length | DB ID | Description |
|-----------|-------|-------------|--------|------------|---------------------|
| 1 | 39 | 100.0 | 9 | 4 AAB48919 | Generic S |
| 2 | 39 | 100.0 | 9 | 4 AAB48917 | AAB48917 SH2 domai |
| 3 | 39 | 100.0 | 9 | 4 AAB48922 | AAB48922 SH2 domai |
| 4 | 39 | 100.0 | 9 | 5 ABG68582 | ABG68582 Peptide G |
| 5 | 39 | 100.0 | 10 | 4 AAB48923 | AAB48923 SH2 domai |
| 6 | 39 | 100.0 | 10 | 4 AAB48920 | AAB48920 SH2 domai |
| 7 | 39 | 100.0 | 10 | 4 AAB48926 | AAB48926 SH2 domai |
| 8 | 39 | 100.0 | 10 | 4 AAB48921 | AAB48921 SH2 domai |
| 9 | 39 | 100.0 | 10 | 4 AAB48928 | AAB48928 SH2 domai |
| 10 | 39 | 100.0 | 11 | 2 AAU46897 | AAU46897 SH2 domai |
| 11 | 39 | 100.0 | 11 | 2 AAU46896 | AAU46896 Non-phosph |
| 12 | 39 | 100.0 | 11 | 5 ABG68419 | ABG68419 G1 peptid |
| 13 | 39 | 100.0 | 11 | 5 ABG68583 | ABG68583 Peptide G |
| 14 | 39 | 100.0 | 26 | 4 AAB48932 | AAB48932 SH2 domai |
| 15 | 39 | 100.0 | 26 | 4 AAB48923 | AAB48923 SH2 domai |
| 16 | 33 | 84.6 | 10 | 4 AAB48925 | AAB48925 SH2 domai |
| 17 | 33 | 84.6 | 10 | 4 AAB48927 | AAB48927 SH2 domai |
| 18 | 33 | 84.6 | 475 | 6 ABU40815 | ABU40815 Protein e |
| 19 | 33 | 84.6 | 586 | 7 ADC39248 | ADC39248 Novel hum |
| 20 | 33 | 84.6 | 592 | 7 ADE86367 | ADE86367 Human PTP |
| 21 | 33 | 84.6 | 593 | 2 AAR52951 | AAR52951 Human pro |
| 22 | 33 | 84.6 | 593 | 2 AAR99313 | AAR99313 Human SH- |
| 23 | 33 | 84.6 | 593 | 2 AAY13476 | AAY13476 Peptide S |
| 24 | 33 | 84.6 | 593 | 4 AAB59218 | AAB59218 SHP-2 act |
| 25 | 33 | 84.6 | 593 | 4 AAB59223 | AAB59223 SHP-2 act |

ALIGNMENTS

RESULT 1
AAB48919
ID AAB48919 standard; peptide; 9 AA.
XX AAB48919;

DT 16-MAR-2001 (first entry)

DE Generic SH2 domain cyclic peptide inhibitor, SEQ ID NO:3.

XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytosolic; tumour; breast cancer; cyclic.

CS Synthetic.

EH Key Location/Qualifiers

FT Modified-site 1..9

FT /note= "The nitrogen atoms of the N-terminus and the C-terminal amide are joined via a bridging moiety, thereby cyclising the peptide"

FT Misc-difference 1

FT /note= "Any naturally or non-naturally occurring amino acid except Glu"

FT Modified-site 9 /note= "C-terminal amide"

FN WO2000073326-A2.

PD 07-DEC-2000.

PP 02-JUN-2000; 2000WO-US015201.

PR 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PL Roller PP, Long Y, Lung FT, King CR, Yang D;

XX WPI; 2001-137633/14.

XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src homology 2 domain binding to target protein, useful for preventing cancer, especially breast cancer.

PS Disclosure; Page 5; 26pp; English.

XX The invention relates to redox-stable, non-phosphorylated cyclic peptides which bind to Src homology 2 (SH2) domains, preventing them from binding

to phosphotyrosine (pTyr)-containing regions of target proteins. The cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4-Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-amino adipic acid (Aad, referred to as Adi in the specification); and Xaa3 is either Aad or Glu. Optionally, there is a conservative or neutral amino acid substitution at either or both of Leu2 and Gly7, and optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified. The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene, which links the nitrogen atom of the N terminus to the nitrogen atom of the C-terminal amide. The peptides are characterised by an in vivo IC-50 of less than 4.0 micromolar when the target protein is Grb2 (growth factor receptor-bound protein 2). On binding Grb2, the peptides have a turn conformation. The peptides, and compositions comprising the peptides, are useful for inhibiting the binding of the SH2 domain to a target protein. They are particularly useful for preventing cancer, especially breast cancer. The present sequence is a generic representation of a cyclic peptide of the invention

XX Sequence 9 AA;

Query Match 100.0%; Score 39; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.4e+06;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9

DB 1 XLYENVGMX 9

RESULT 2

AAB48917

ID AAB48917 standard; peptide; 9 AA.

XX AAB48917;

DT 16-MAR-2001 (first entry)

DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:1.

XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;

KW cytostatic; tumour; breast cancer; cyclic.

XX Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1..9

/note= "The nitrogen atoms of the N-terminus and the C-terminal amide are joined via a bridging moiety, thereby cyclising the peptide"

FT Modified-site 1

FT Modified-site 9 /note= "Gamma-carboxyglutamic acid"

FT Modified-site 9 /note= "C-terminal amide"

XX WO2000073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Roller PP, Long Y, Lung FT, King CR, Yang D;

XX WPI; 2001-137633/14.

XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src

FT homology 2 domain binding to target protein, useful for preventing

PT cancer, especially breast cancer.

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Claim 1; Page 21; 26pp; English.

The invention relates to redox-stable, non-phosphorylated cyclic peptides which bind to Src homology 2 (SH2) domains, preventing them from binding to phosphotyrosine (pTyr)-containing regions of target proteins. The cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4-Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-amino adipic acid (Aad, referred to as Adi in the specification); and Xaa3 is either Aad or Glu. Optionally, there is a conservative or neutral amino acid substitution at either or both of Leu2 and Gly7, and optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified. The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene, which links the nitrogen atom of the N terminus to the nitrogen atom of the C-terminal amide. The peptides are characterised by an in vivo IC-50 of less than 4.0 micromolar when the target protein is Grb2 (growth factor receptor-bound protein 2). On binding Grb2, the peptides have a turn conformation. The peptides, and compositions comprising the peptides, are useful for inhibiting the binding of the SH2 domain to a target protein. They are particularly useful for preventing cancer, especially breast cancer. The present sequence represents a cyclic peptide of the invention

XX Sequence 9 AA;

Query Match 100.0%; Score 39; DB 4; Length 9;

Best Local Similarity 88.9%; Pred. No. 1.4e+06;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9

DB 1 XLYENVGMX 9

RESULT 3

AAB48922

ID AAB48922 standard; peptide; 9 AA.

XX AAB48922;

DT 16-MAR-2001 (first entry)

DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:7.

KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;

KW cytostatic; tumour; breast cancer; linear precursor.

XX Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1

/note= "Gamma-carboxyglutamic acid; the nitrogen atom of the N-terminus is joined to a ClCH2C(O) moiety"

FT Modified-site 9

/note= "The carbon atom of the C-terminus is joined to a C(CH2SH)C(O)NH2 moiety"

XX WO2000073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Roller PP, Long Y, Lung FT, King CR, Yang D;

XX WPI; 2001-137633/14.

PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
XX cancer, especially breast cancer.
PS Example 1; Page 13; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pY)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC amino adipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 39; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.4e+06;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 XLVENVGMY 9
DB :|||||:
1 XLVENVGMY 9
RESULT 4
ID ABG68582 standard; peptide; 9 AA.
XX
AC ABG68582;
XX
XX 07-OCT-2002 (first entry)
DT Peptide GLTE #1.
DE
XX Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; GI;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
XX WO200236142-A2.
PN
XX 10-MAY-2002.
PD
XX 05-NOV-2001; 2001WO-US047400.
PE
XX 03-NOV-2000; 2000US-0245755P.
PR (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
XX Krag DN, Pero SC, Oligino L;
XX WPI; 2002-547451/58.
DR
XX Treatment or prophylaxis of a subject having a disorder characterized by

PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Fig 9B; 186pp; English.
XX
CC The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC the nucleic acid encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phage display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal
CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC placenta, colon, ovary, testes and lung. The present sequence is a GI
CC peptide (not defined) or derivative which is used to illustrate the
CC possible structures of cyclic Grb7 antagonists
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 39; DB 5; Length 9;
Best Local Similarity 77.8%; Pred. No. 1.4e+06;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 XLVENVGMY 9
DB :|||||:
1 ELVENVGMY 9
RESULT 5
ID AAB48923 standard; peptide; 10 AA.
XX
AC AAB48923;
XX
XX 16-MAR-2001 (first entry)
DT
DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:8.
XX
XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; cyclic.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1..10
FT /note= "The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety, thereby
FT cyclising the peptide"
FT Modified-site 1
FT /label= Aad
FT Modified-site 10
FT /note= "C-terminal amide"
XX
XX WO200073326-A2.
PN
XX 07-DEC-2000.
PD
XX 02-JUN-2000; 2000WO-US015201.
PF
XX 02-JUN-1999; 99US-0137187P.
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Roller PP, Long Y, Lung FT, King CR, Yang D;
XX

```

DR  WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
XX Example 2; Page 13; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC amino adipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
XX peptide of the invention
XX
XX Sequence 10 AA;
XX
XX Query Match 100.0%; Score 39; DB 4; Length 10;
XX Best Local Similarity 88.9%; Pred. No. 0.09;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 XLYENVGMX 9
XX | | | | |
XX Db 1 XLYENVGMX 9
XX
XX RESULT 6
XX AAB48920
XX ID AAB48920 standard; peptide; 10 AA.
XX AC AAB48920;
XX
XX DT 16-MAR-2001 (first entry)
XX
XX DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:4.
XX
XX KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytostatic; tumour; breast cancer; cyclic.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1..10
XX /note= "The nitrogen atoms of the N-terminus and the C-
XX terminal amide are joined via a bridging moiety C(O)-CH2-
XX S-CH2-CHC(O)NH2, thereby cyclising the peptide"
XX
XX FT Modified-site 1
XX /note= "Gamma-carboxyglutamic acid"
XX
XX FT Modified-site 10
XX /note= "C-terminal amide"
XX
XX PN WO200073326-A2.
XX
XX PD 07-DEC-2000.
XX
XX PF 02-JUN-2000; 2000WO-US015201.
XX
XX PR 02-JUN-1999; 99US-0137187P.

```

```

XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
XX WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
XX Example 1; Page 12; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC amino adipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
XX peptide of the invention
XX
XX Sequence 10 AA;
XX
XX Query Match 100.0%; Score 39; DB 4; Length 10;
XX Best Local Similarity 88.9%; Pred. No. 0.09;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 XLYENVGMX 9
XX | | | | |
XX Db 1 XLYENVGMX 9
XX
XX RESULT 7
XX AAB48926
XX ID AAB48926 standard; peptide; 10 AA.
XX AC AAB48926;
XX
XX DT 16-MAR-2001 (first entry)
XX
XX DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:11.
XX
XX KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytostatic; tumour; breast cancer; linear precursor.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Modified-site 10
XX /label= Nle
XX /note= "C-terminal amide, joined to a solid matrix"
XX
XX PN WO200073326-A2.
XX
XX PD 07-DEC-2000.
XX
XX PF 02-JUN-2000; 2000WO-US015201.
XX
XX PR 02-JUN-1999; 99US-0137187P.

```

```

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Roller PP, Long Y, Lung FT, King CR, Yang D;
XX WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
XX Example 4; Page 14; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pYr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu3-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
XX Sequence 10 AA;
SQ
Query Match 100.0%; Score 39; DB 4; Length 10;
Best Local Similarity 77.8%; Pred.No. 0.09;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 XLYENVGMX 9
DB :|||||:
1 ELYENVGMY 9

RESULT 8
AAB48921
ID AAB48921 standard; peptide; 10 AA.
XX
XX AAB48921;
XX
XX 16-MAR-2001 (first entry)
XX
XX SH2 domain peptide inhibitor linear precursor, SEQ ID NO:5.
XX
XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytostatic; tumour; breast cancer; linear precursor.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "Gamma-carboxyglutamic acid"
FT
XX WO200073326-A2.
XX
XX 07-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015201.
XX
XX 02-JUN-1999; 99US-0137187P.
XX

```

```

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Roller PP, Long Y, Lung FT, King CR, Yang D;
XX WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
XX Example 1; Page 12; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pYr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
XX Sequence 10 AA;
SQ
Query Match 100.0%; Score 39; DB 4; Length 10;
Best Local Similarity 88.9%; Pred.No. 0.09;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 XLYENVGMX 9
DB :|||||:
1 XLYENVGMY 9

RESULT 9
AAB48928
ID AAB48928 standard; peptide; 10 AA.
XX
XX AAB48928;
XX
XX 16-MAR-2001 (first entry)
XX
XX SH2 domain peptide inhibitor linear precursor, SEQ ID NO:14.
XX
XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytostatic; tumour; breast cancer; linear precursor.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 10
FT /label= Aad
FT /note= "C-terminal amide, joined to a solid matrix"
FT
XX WO200073326-A2.
XX
XX 07-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015201.
XX
XX 02-JUN-1999; 99US-0137187P.
XX

```

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Roller PP, Long Y, Lung FT, King CR, Yang D;
 PI WPI; 1998-110340/10.
 XX
 XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 PT homology 2 domain binding to target protein, useful for preventing
 PT cancer, especially breast cancer.
 XX
 XX Example 5; Page 15; 26pp; English.
 XX
 XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to Src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC amino adipic acid (Aad), referred to as Adi in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a linear
 CC precursor of a peptide of the invention
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 100.0%; Score 39; DB 4; Length 10;
 Best Local Similarity 77.8%; Pred. No. 0.09;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVGMX 9
 Db :|||||:
 1-ELYENVGMV 9
 RESULT 10
 AAW46897
 ID AAW46897 standard; peptide; 11 AA.
 XX
 XX AAW46897;
 AC
 XX 19-JUN-1998 (first entry)
 DT
 XX G1C-S peptide.
 DE
 XX SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
 KW signal transduction protein; non-phosphorylated; inhibition; treatment;
 KW hyper-proliferative disease; human cancer.
 XX
 XX Unidentified.
 OS
 XX WO9802176-A1.
 PN
 XX 22-JAN-1998.
 PD
 XX 16-JUL-1997; 97WO-US012501.
 PF
 XX 16-JUL-1996; 96US-0021858P.
 PR
 XX (GEOU) UNIV GEORGETOWN.
 PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX
 XX King CR, Sastry L, Krag D, Oligino L;
 PI WPI; 1998-110340/10.
 XX
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein -at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.

XX WPI; 1998-110340/10.
 DR
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein - at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.
 XX
 XX Disclosure; Page 18; 39pp; English.
 PS
 XX The present sequence represents a peptide designated G1C-S. This peptide
 CC is essentially the same as a non-phosphorylated peptide, G1, that is
 CC capable of binding to the src homology 2 (SH2) domain of Grb2, except
 CC that the terminal Cys residues of G1 are replaced with Ser residues. Grb2
 CC is a signal transduction protein. The binding affinity of the present
 CC peptide with Grb2 was tested, and it was demonstrated that the disulphide
 CC bond of G1 may be important. The G1 peptide binds to the SH2 domain of
 CC Grb2 with affinity similar to, or greater than, that of a SHC
 CC phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue
 CC that has not been modified by phosphate or similar charged group. The G1
 CC peptide is used to inhibit a signal transduction process that involves
 CC binding of a phosphorylated protein or peptide to the SH2 domain of a
 CC signal transduction protein, particularly Grb2. It is used specifically
 CC for treatment of hyper-proliferative diseases, especially human cancer
 XX
 XX Sequence 11 AA;
 SQ
 Query Match 100.0%; Score 39; DB 2; Length 11;
 Best Local Similarity 77.8%; Pred. No. 0.1;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVGMX 9
 Db :|||||:
 2 ELYENVGMV 10
 RESULT 11
 AAW46896
 ID AAW46896 standard; peptide; 11 AA.
 XX
 XX AAW46896;
 AC
 XX 19-JUN-1998 (first entry)
 DT
 XX Non-phosphorylated peptide which binds to the SH2 domain of Grb2.
 DE
 XX SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
 KW signal transduction protein; non-phosphorylated; inhibition; treatment;
 KW hyper-proliferative disease; human cancer; cyclic.
 XX
 XX Unidentified.
 OS
 XX Disulfide-bond 1..11
 FH
 XX WO9802176-A1.
 PN
 XX 22-JAN-1998.
 PD
 XX 16-JUL-1997; 97WO-US012501.
 PF
 XX 16-JUL-1996; 96US-0021858P.
 PR
 XX (GEOU) UNIV GEORGETOWN.
 PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX
 XX King CR, Sastry L, Krag D, Oligino L;
 PI WPI; 1998-110340/10.
 XX
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein -at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.

PS Claim 9; Page 17; 39pp; English.

CC The present sequence represents non-phosphorylated peptide, G1, that is
 CC capable of binding to the src homology 2 (SH2) domain of Grb2. Grb2 is a
 CC signal transduction protein. The G1 peptide binds to the SH2 domain of
 CC Grb2 with affinity similar to, or greater than, that of a SHC
 CC phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue
 CC that has not been modified by phosphate or similar charged group. The G1
 CC peptide is used to inhibit a signal transduction process that involves
 CC binding of a phosphorylated protein or peptide to the SH2 domain of a
 CC signal transduction protein, particularly Grb2. It is used specifically
 CC for treatment of hyper-proliferative diseases, especially human cancer
 XX

XX Sequence 11 AA;

Query Match 100.0%; Score 39; DB 2; Length 11;
 Best Local Similarity 77.8%; Pred. No. 0.1;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
 :|||||:
 Db 2 ELYENVGY 10

RESULT 12

ABG68419

ID ABG68419 standard; peptide; 11 AA.

AC ABG68419;

XX

DT 07-OCT-2002 (first entry)

DE G1 peptide.

KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
 KW cytosolic; cancer; phage display; tumour; metastasis; breast cancer;
 KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
 KW breast disorder; oesophageal disorder; pancreatic disorder; GI;
 KW prostate disorder; small intestine disorder; placental disorder;
 KW colon disorder; ovary disorder; testicular disorder; lung disorder.
 XX Synthetic.
 OS
 XX WO200236142-A2.
 XX 10-MAY-2002.
 XX 05-NOV-2001; 2001WO-US047400.
 XX 03-NOV-2000; 2000US-0245755P.
 XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX Krag DN, Pero SC, Oligino L;
 XX WPI; 2002-547451/58.
 XX Treatment or prophylaxis of a subject having a disorder characterized by
 XX abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
 XX a non-phosphorylated peptide to a subject in need of the treatment.
 XX Disclosure; Page 102; 186pp; English.

CC The invention relates to treatment or prophylaxis (M1) of a subject
 CC having a disorder characterised by abnormal interaction of Grb7 (Growth
 CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
 CC administering to a subject in need of the treatment, a non-phosphorylated
 CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
 CC Asn) or its functional equivalent, in an amount effective to inhibit the
 CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
 CC nucleic acids encoding the antagonists, an expression vector comprising
 CC the nucleic acid, a host cell transformed or transfected with the vector,
 CC screening (M2) a molecular library to identify a compound that inhibits
 CC interaction between Grb7 and a peptide antagonist and a phage display
 CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
 CC treatment of a subject having a disorder characterised by abnormal

CC interaction between Grb7 and a peptide antagonist and a phage display
 CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
 CC treatment of a subject having a disorder characterised by abnormal
 CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
 CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
 CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
 CC placenta, colon, ovary, testes and lung. The present sequence is a G1
 CC peptide (not defined) or derivative which is used to illustrate the
 XX possible structures of cyclic Grb7 antagonists

XX Sequence 11 AA;

Query Match 100.0%; Score 39; DB 5; Length 11;
 Best Local Similarity 77.8%; Pred. No. 0.1;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
 :|||||:
 Db 2 ELYENVGY 10

RESULT 13

ABG68583

ID ABG68583 standard; peptide; 11 AA.

XX

AC ABG68583;

XX

DT 07-OCT-2002 (first entry)

DE Peptide G1TE #2.

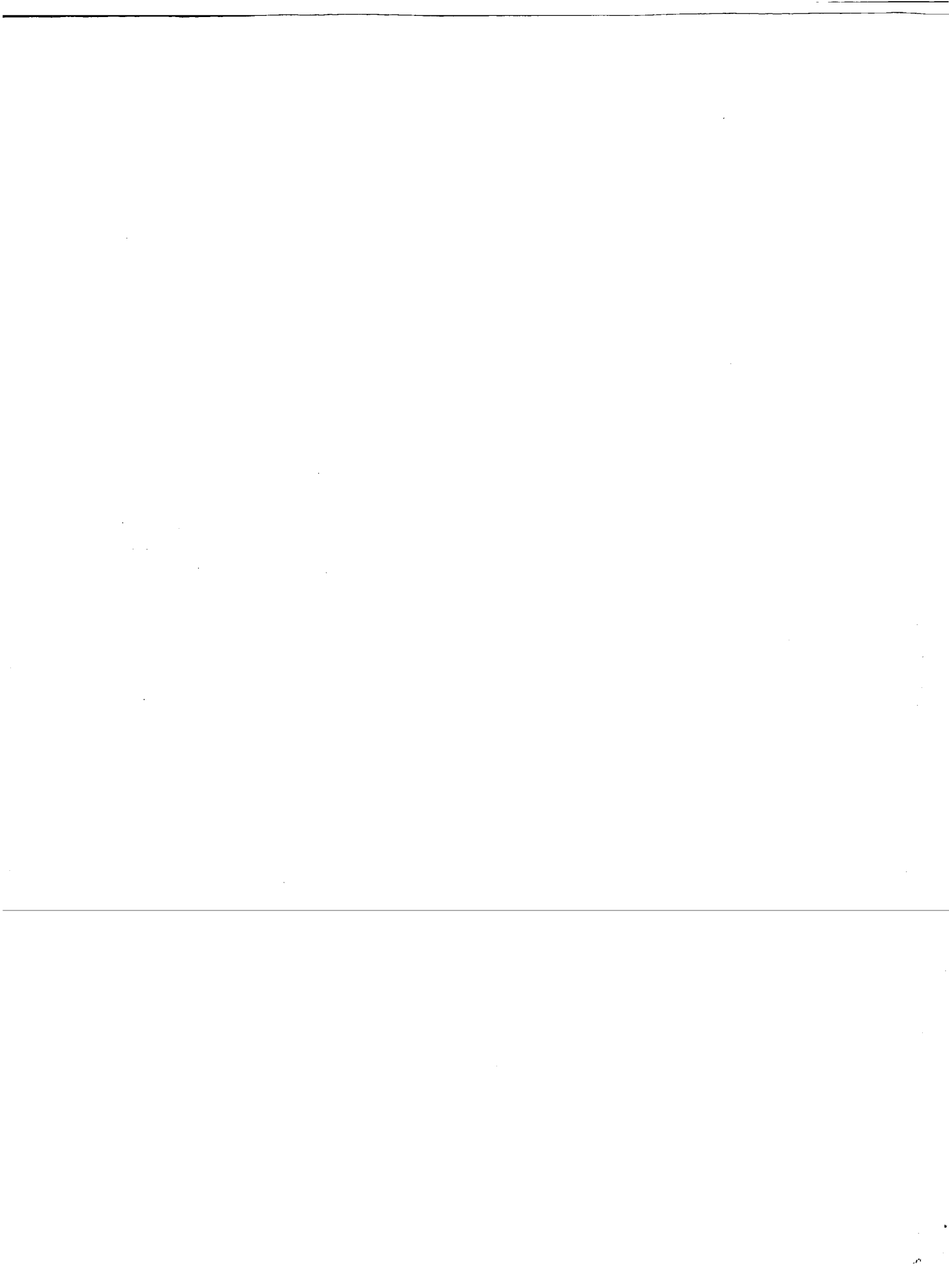
KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
 KW cytosolic; cancer; phage display; tumour; metastasis; breast cancer;
 KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
 KW breast disorder; oesophageal disorder; pancreatic disorder; GI;
 KW prostate disorder; small intestine disorder; placental disorder;
 KW colon disorder; ovary disorder; testicular disorder; lung disorder.
 XX Synthetic.
 OS
 XX WO200236142-A2.
 XX 10-MAY-2002.
 XX 05-NOV-2001; 2001WO-US047400.
 XX 03-NOV-2000; 2000US-0245755P.
 XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX Krag DN, Pero SC, Oligino L;
 XX WPI; 2002-547451/58.
 XX Treatment or prophylaxis of a subject having a disorder characterized by
 XX abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
 XX a non-phosphorylated peptide to a subject in need of the treatment.
 XX Disclosure; Fig 9C; 186pp; English.

CC The invention relates to treatment or prophylaxis (M1) of a subject
 CC having a disorder characterised by abnormal interaction of Grb7 (Growth
 CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
 CC administering to a subject in need of the treatment, a non-phosphorylated
 CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
 CC Asn) or its functional equivalent, in an amount effective to inhibit the
 CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
 CC nucleic acids encoding the antagonists, an expression vector comprising
 CC the nucleic acid, a host cell transformed or transfected with the vector,
 CC screening (M2) a molecular library to identify a compound that inhibits
 CC interaction between Grb7 and a peptide antagonist and a phage display
 CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
 CC treatment of a subject having a disorder characterised by abnormal

CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 39; DB 4; Length 26;
Best Local Similarity 88.9%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 XLYENVGMX 9
DB 1 XLYENVGMY 9

Search completed: July 15, 2004, 07:28:50
Job time : 49 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:26:37 ; Search time 14.5 Seconds
(without alignments)
32.044 Million cell updates/sec

Title: SEQ1MOD
Perfect score: 39
Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pdp:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pdp:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pdp:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pdp:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pdp:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----------------------|--------------------|
| 1 | 33 | 84.6 | 485 | US-09-543-681A-4935 | Sequence 4935, Ap |
| 2 | 33 | 84.6 | 566 | US-08-272-255-8 | Sequence 8, Appli |
| 3 | 33 | 84.6 | 566 | PCT-US95-08565-8 | Sequence 8, Appli |
| 4 | 33 | 84.6 | 593 | US-08-202-389-12 | Sequence 12, Appli |
| 5 | 33 | 84.6 | 593 | US-08-018-129-5 | Sequence 5, Appli |
| 6 | 33 | 84.6 | 593 | US-08-448-250-5 | Sequence 5, Appli |
| 7 | 33 | 84.6 | 593 | US-09-282-237-5 | Sequence 5, Appli |
| 8 | 32 | 82.1 | 214 | US-09-489-039A-12637 | Sequence 12637, A |
| 9 | 31 | 79.5 | 325 | US-09-134-001C-3551 | Sequence 3551, Ap |
| 10 | 31 | 79.5 | 335 | US-09-134-000C-3814 | Sequence 3814, Ap |
| 11 | 31 | 79.5 | 919 | US-08-788-674-4 | Sequence 4, Appli |
| 12 | 30 | 76.9 | 19 | US-09-376-343-3 | Sequence 3, Appli |
| 13 | 30 | 76.9 | 20 | US-08-480-190-38 | Sequence 38, Appli |
| 14 | 30 | 76.9 | 20 | US-08-475-399A-38 | Sequence 38, Appli |
| 15 | 30 | 76.9 | 20 | PCT-US93-07545-38 | Sequence 38, Appli |
| 16 | 30 | 76.9 | 20 | US-09-107-532A-6346 | Sequence 6346, Ap |
| 17 | 30 | 76.9 | 141 | US-09-540-236-3016 | Sequence 3016, Ap |
| 18 | 30 | 76.9 | 180 | US-09-003-287-8 | Sequence 6, Appli |
| 19 | 30 | 76.9 | 244 | US-09-003-287-8 | Sequence 8, Appli |
| 20 | 30 | 76.9 | 244 | US-09-518-988-2 | Sequence 2, Appli |
| 21 | 30 | 76.9 | 362 | US-09-080-897-6 | Sequence 6, Appli |
| 22 | 30 | 76.9 | 362 | US-09-323-735-6 | Sequence 6, Appli |
| 23 | 30 | 76.9 | 459 | US-09-543-681A-6287 | Sequence 6287, Ap |
| 24 | 30 | 76.9 | 1250 | US-08-938-291A-9 | Sequence 9, Appli |
| 25 | 30 | 76.9 | 1250 | US-09-589-619-9 | Sequence 9, Appli |
| 26 | 30 | 76.9 | 1250 | US-08-146-145-6 | Sequence 6, Appli |
| 27 | 29 | 74.4 | 9 | | |

ALIGNMENTS

RESULT 1
US-09-543-681A-4935
; Sequence 4935, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4935
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4935

Query Match 84.6%; Score 33; DB 4; Length 485;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Cy 1 XLYENVGMX 9
Db 457 TLIESIGMA 465

RESULT 2
US-08-272-255-8
; Sequence 8, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

Sequence 189, App
Sequence 8937, Ap
Sequence 32806, A
Sequence 134, App
Sequence 12406, A
Sequence 4, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 2994, Ap
Sequence 6700, Ap
Sequence 11, Appli
Sequence 6, Appli
Patent No. 5422248
Sequence 4, Appli
Sequence 8, Appli
Sequence 25011, A
Sequence 6895, Ap

28 29 74.4 127 3 US-08-467-023-189
29 29 74.4 204 4 US-09-489-039A-8937
30 29 74.4 475 4 US-09-252-991A-32806
31 29 74.4 514 3 US-08-467-023-134
32 29 74.4 531 4 US-09-489-039A-12406
33 29 74.4 574 3 US-09-385-028-4
34 29 74.4 574 4 US-09-726-614-4
35 29 74.4 574 4 US-09-385-040-4
36 29 74.4 602 2 US-08-419-852-6
37 29 74.4 607 4 US-09-134-001C-2994
38 29 74.4 617 4 US-09-328-352-6700
39 29 74.4 698 3 US-08-941-445A-11
40 29 74.4 771 1 US-07-923-976-6
41 29 74.4 783 6 5422248-2
42 29 74.4 836 1 US-07-923-976-4
43 29 74.4 863 1 US-07-923-976-8
44 29 74.4 900 4 US-09-252-991A-25011
45 28 71.8 76 4 US-09-621-976-6895

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 08-JUL-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: UPN-1795
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-272-255-8

Query Match 84.6%; Score 33; DB 2; Length 566;
Best Local Similarity 55.6%; Pred. No. 78;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 87 RLYDNVGLY 95

RESULT 3
PCT-US95-08565-8
SEQUENCE 8, Application PC/TUS9508565
GENERAL INFORMATION:
APPLICANT: Cashmore, Anthony R.
APPLICANT: Ahmad, Margaret
APPLICANT: Lin, Chentao
TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: PCT/US95/08565
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/272,255
FILING DATE: 08-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: UPN-1795
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
PCT-US95-08565-8

Query Match 84.6%; Score 33; DB 5; Length 566;
Best Local Similarity 55.6%; Pred. No. 78;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 87 RLYDNVGLY 95

RESULT 4
US-08-202-389-12
SEQUENCE 12, Application US/08202389
PATENT NO. 5536636
GENERAL INFORMATION:
APPLICANT: Freeman Jr., Robert M.
APPLICANT: Plutsky, Jorge
APPLICANT: Neel, Benjamin G.
APPLICANT: Rosenberg, Robert D.
TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,389
FILING DATE: 28-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,926
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/829,141
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/721,112
FILING DATE: 26-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: BIH92-05MA
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 593 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-202-389-12

Query Match 84.8%; Score 33; DB 1; Length 593;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 578 RYENVGLM 586

```
RESULT 5
US-08-018-129-5
; Sequence 5, Application US/08018129
; Patent No. 5589375
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP 1D: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/018,129
; FILING DATE: 19930216
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-018-129-5
Query Match 84.6%; Score 33; DB 1; Length 593;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 578 RYENVGLM 586

RESULT 6
US-08-448-250-5
; Sequence 5, Application US/08448250
; Patent No. 5951251
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP 1D: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,250
; FILING DATE: 19930216
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-448-250-5
Query Match 84.6%; Score 33; DB 1; Length 593;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 578 RYENVGLM 586

RESULT 7
US-09-282-257-5
; Sequence 5, Application US/09282257
; Patent No. 6548641
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP 1D: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/282,257
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/018,129
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,250
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/018,129
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-448-250-5
Query Match 84.6%; Score 33; DB 2; Length 593;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 578 RYENVGLM 586

RESULT 7
US-09-282-257-5
; Sequence 5, Application US/09282257
; Patent No. 6548641
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP 1D: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/282,257
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/018,129
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-282-257-5

Query Match      84.6%; Score 33; DB 4; Length 593;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 XLYENVGMX 9
Db      578 RVYENVGLM 586

RESULT 8
US-09-489-039A-12637
; Sequence 12637, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12637
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12637

Query Match      82.1%; Score 32; DB 4; Length 214;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 XLYENVGMX 9
Db      148 SLYENVGMA 156

RESULT 9
US-09-134-001C-3551
; Sequence 3551, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3551
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3551

Query Match      79.5%; Score 31; DB 4; Length 325;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 XLYENVGMX 9
Db      167 QYVESIGMD 175

RESULT 10
US-09-134-000C-3814
; Sequence 3814, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3814
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3814

Query Match      79.5%; Score 31; DB 4; Length 335;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 XLYENVGMX 9
Db      53 LLYXNTGMT 61

RESULT 11
US-08-788-674-4
; Sequence 4, Application US/08788674
; Patent No. 5922315
; GENERAL INFORMATION:
; APPLICANT: Roy, Soumitra
; TITLE OF INVENTION: Adenoviruses Having Altered
; TITLE OF INVENTION: Hexon Proteins
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carelia, Byrne, Bain,
; ADDRESSEE: Gilfillan, Cecchi, Stewart &
; ADDRESSEE: Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,674
; FILING DATE: 24-JAN-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 271010-363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 919 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

```

```
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: predicted hexon protein sequence
; NAME/KEY: for human adenovirus 12
US-08-788-674-4

Query Match      79.5%; Score 31; DB 2; Length 919;
Best Local Similarity 55.6%; Pred. No. 3.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 439 FLYSNVGLY 447

RESULT 12
US-09-376-343-3
; Sequence 3, Application US/09376343
; Patent No. 6506592
; GENERAL INFORMATION:
; APPLICANT: Blum, Paul H.
; TITLE OF INVENTION: Hyperthermophilic Alpha-Glucosidase Gene and Its Use
; FILE REFERENCE: N1231-200
; CURRENT APPLICATION NUMBER: US/09/376,343
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/096,860
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-09-376-343-3

Query Match      76.9%; Score 30; DB 4; Length 19;
Best Local Similarity 44.4%; Pred. No. 7.3;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 5 KYENLVGY 13

RESULT 13
US-08-480-190-38
; Sequence 38, Application US/08480190
; Patent No. 5827516
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,190
; FILING DATE:
; CLASSIFICATION: 424

; INFORMATION FOR SEQ ID NO: 424
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-8906
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-480-190-38

Query Match      76.9%; Score 30; DB 2; Length 20;
Best Local Similarity 71.4%; Pred. No. 7.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7
Db 2 TLYQNVG 8

RESULT 14
US-08-488-379-38
; Sequence 38, Application US/08488379
; Patent No. 5880103
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,379
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 38:
```

Search completed: July 15, 2004, 07:31:19
Job time : 15.5 secs

; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-488-379-38

Query Match 76.9%; Score 30; DB 2; Length 20;
Best Local Similarity 71.4%; Pred. No. 7.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Caps 0;

Qy 1 XLYENVG 7
Db 2 TLYQNVG 8

RESULT 15
US-08-475-399A-38
; Sequence 38, Application US/08475399A
; Patent No. 6509033
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicz, Roman M.
; APPLICANT: Vignali, Dario A.A.
; APPLICANT: Hedley, Mary L.
; APPLICANT: Stern, Lawrence J.
; APPLICANT: Strominger, Jack L.
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 276
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,399A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: 15-JUN-1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00246/168003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-507
; TELEFAX: 617/542-890
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-475-399A-38

Query Match 76.9%; Score 30; DB 4; Length 20;
Best Local Similarity 71.4%; Pred. No. 7.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Caps 0;

Qy 1 XLYENVG 7
Db 2 TLYQNVG 8

; LOCATION: (1)..(9)
; OTHER INFORMATION: Xaa (Gla) and Tyr at position 9 are bridged together, making this
; OTHER INFORMATION: peptide cyclic
US-09-998-350-1

Query Match 100.0%; Score 39; DB 10; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.2e+06;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMV 9

RESULT 2

US-09-998-350-3
; Sequence 3, Application US/09998350
; Publication No. US20030078368A1

GENERAL INFORMATION:

APPLICANT: Roller, Peter P

APPLICANT: Long, Ya-Qiu

APPLICANT: Lung, Feng-Di T

APPLICANT: King, Richter C

APPLICANT: Yang, Dajun

TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2

TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND

FILE REFERENCE: 214683

CURRENT APPLICATION NUMBER: US/09/998,350

PRIOR FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: PCT/US00/15201

PRIOR FILING DATE: 2000-06-02

PRIOR APPLICATION NUMBER: 60/137,187

PRIOR FILING DATE: 1999-06-02

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(1)

OTHER INFORMATION: Xaa is any amino acid other than Glu

FEATURE:

NAME/KEY: misc_feature

LOCATION: (9)..(9)

OTHER INFORMATION: Tyr at position 9 is an amide, i.e., C(O)NH

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(9)

OTHER INFORMATION: Xaa and Tyr at position 9 are bridged together, making this pepti

OTHER INFORMATION: de cyclic

US-09-998-350-3

Query Match 100.0%; Score 39; DB 10; Length 9;

Best Local Similarity 88.9%; Pred. No. 1.2e+06;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMV 9

RESULT 3

US-09-998-350-7

; Sequence 7, Application US/09998350

; Publication No. US20030078368A1

APPLICANT: Roller, Peter P

APPLICANT: Long, Ya-Qiu

; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa has a ClCH2C(O)- group attached
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 has a -C(CH2SH)C(O)NH2 group attached
US-09-998-350-7

Query Match 100.0%; Score 39; DB 10; Length 9;

Best Local Similarity 88.9%; Pred. No. 1.2e+06;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMV 9

RESULT 4

US-09-998-350-4

; Sequence 4, Application US/09998350

; Publication No. US20030078368A1

GENERAL INFORMATION:

APPLICANT: Roller, Peter P

APPLICANT: Long, Ya-Qiu

APPLICANT: Lung, Feng-Di T

APPLICANT: King, Richter C

APPLICANT: Yang, Dajun

TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2

TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND

FILE REFERENCE: 214683

CURRENT APPLICATION NUMBER: US/09/998,350

CURRENT FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: PCT/US00/15201

PRIOR FILING DATE: 2000-06-02

PRIOR APPLICATION NUMBER: 60/137,187

PRIOR FILING DATE: 1999-06-02

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4

LENGTH: 10

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys at position 10 is an amide, i.e., C(O)NH
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Xaa (Gla) and Cys are bridged together, making this peptide cyclic
; OTHER INFORMATION: C
US-09-998-350-4

```

```

Query Match      100.0%; Score 39; DB 10; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMY 9

```

```

RESULT 5
US-09-998-350-5
; Sequence 5, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
US-09-998-350-5

```

```

Query Match      100.0%; Score 39; DB 10; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMY 9

```

```

RESULT 6
US-09-998-350-6
; Sequence 6, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
US-09-998-350-6

```

```

Query Match      100.0%; Score 39; DB 10; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMY 9

```

```

RESULT 7
US-09-998-350-7
; Sequence 7, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2

```

```

; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND N
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla(OtBu)2, which is di- tert-butoxy-gamma-carboxy-L-glutam
; OTHER INFORMATION: ic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Tyr at position 3 is modified to Tyr(tBu), which is tert-butyl-ty
; OTHER INFORMATION: rosine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is is trytyl-asp
; OTHER INFORMATION: argine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(tBu), which is tert-butyl-ty
; OTHER INFORMATION: rosine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys at position 10 is modified to Cys(Trt), which is trytyl-cyste
; OTHER INFORMATION: ine, and Cys(Trt) is connected to a resin
US-09-998-350-6

```

```

Query Match      100.0%; Score 39; DB 10; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMY 9

```

```

RESULT 7
US-09-998-350-8
; Sequence 8, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2

```

```

; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Adi, which is alpha-amino-adipic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa has a CH2CO- group attached
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Xaa (Adi) and Cys are bridged together, making this peptide cycli
US-09-998-350-8

Query Match 100.0%; Score 39; DB 10; Length 10;
Best Local Similarity 88.98; Pred. No. 0.24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMY 9

RESULT 8
US-09-998-350-11
; Sequence 11, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Glu at position 1 is modified to Glu(OtBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Tyr at position 3 is modified to Tyr(OtBu), which is tert-butoxy-
; OTHER INFORMATION: tyrosine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trityl-aspara
; OTHER INFORMATION: gine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy-
; OTHER INFORMATION: tyrosine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa = Nle, which is norleucine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa is an amide and is attached to a resin
US-09-998-350-11

Query Match 100.0%; Score 39; DB 10; Length 10;
Best Local Similarity 77.8%; Pred. No. 0.24;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 1 ELXENVGMY 9

RESULT 9
US-09-998-350-14
; Sequence 14, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic

```

; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Glu at position 1 is modified to Glu(OcBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Glu at position 4 is modified to Glu(OcBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trytyl-aspara-
; OTHER INFORMATION: Gine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OcBu), which is tert-butoxy-
; OTHER INFORMATION: tyrosine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa = Adi(OAl), which is allyloxy-alpha-amino-adipic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa is an amide, i.e., C(O)NH
US-09-998-350-14

Query Match 100.0%; Score 39; DB 10; Length 10;
Best Local Similarity 77.8%; Pred. No. 0.24; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0;

Qy 1 XLYENVGMX 9
:|||||:
Db 1 ELYENVGMY 9

RESULT 10
US-10-013-815-32
; Sequence 32, Application US/10013815
; Publication No. US20030105000A1
; GENERAL INFORMATION:
; APPLICANT: Pero, Stephanie
; APPLICANT: Krag, David
; APPLICANT: Oligino, Lyn
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING GRE7
; FILE REFERENCE: V0139/7048 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/10/013,815
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 60/245,755
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030105000A1-phosphorylated peptide with YEN motif
US-10-013-815-32

Query Match 100.0%; Score 39; DB 14; Length 11;
Best Local Similarity 77.8%; Pred. No. 0.27; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0;

Qy 1 XLYENVGMX 9
:|||||:
Db 2 ELYENVGMY 10

RESULT 11
US-09-998-350-19

; Sequence 18, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND N
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
US-09-998-350-18

Query Match 100.0%; Score 39; DB 10; Length 26;
Best Local Similarity 88.9%; Pred. No. 0.68; 0; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0;

Qy 1 XLYENVGMX 9
:|||||:
Db 1 XLYENVGMY 9

RESULT 12
US-09-998-350-19
; Sequence 19, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND N
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid

```
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa (Gla) has a CH2CO- group attached
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
US-09-998-350-19

Query Match      100.0%; Score 39; DB 10; Length 26;
Best Local Similarity 88.9%; Pred. No. 0.68;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 1 XLYENVGMY 9

RESULT 13
US-10-437-963-168439
; Sequence 168439, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168439
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(134)
; OTHER INFORMATION: unsure at all Xaa locations
;
; OTHER INFORMATION: Clone ID: PAT_MRT4530_66953C.1.pep
US-10-437-963-168439

Query Match      87.2%; Score 34; DB 16; Length 134;
Best Local Similarity 55.6%; Pred. No. 42;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
Db 103 DIVENMGK 111

RESULT 14
US-09-998-350-10
; Sequence 10, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683

Query Match      84.6%; Score 33; DB 10; Length 10;
Best Local Similarity 85.7%; Pred. No. 4.1;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVG 7
Db 1 ELYENVG 7

RESULT 15
US-09-998-350-12
; Sequence 12, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa = Nle, which is norleucine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(10)
; OTHER INFORMATION: Glu at position 1 and Cys are bridged together, making this pepti
; OTHER INFORMATION: de cyclic
US-09-998-350-12

Query Match      84.6%; Score 33; DB 10; Length 10;
Best Local Similarity 85.7%; Pred. No. 4.1;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 XLYENVG 7
       :||||
Db      1 ELYENVG 7

Search completed: July 15, 2004, 07:32:50
Job time : 41 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:23:22 ; Search time 11.5 Seconds
(without alignments)
75.280 Million cell updates/sec

Title: SEQ1MOD

Perfect score: 39

Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 34 | 87.2 | 688 | T33708 | hypothetical prote |
| 2 | 33 | 84.6 | 565 | S67298 | deoxyribodipyrimid |
| 3 | 33 | 84.6 | 593 | JN0805 | protein-tyrosine-p |
| 4 | 33 | 84.6 | 593 | JC5167 | protein-tyrosine-p |
| 5 | 33 | 84.6 | 595 | A55651 | protein-tyrosine-p |
| 6 | 33 | 84.6 | 597 | A53593 | protein-tyrosine-p |
| 7 | 33 | 84.6 | 700 | T20550 | hypothetical prote |
| 8 | 32 | 82.1 | 178 | B69944 | hypothetical prote |
| 9 | 32 | 82.1 | 231 | H85138 | hypothetical prote |
| 10 | 32 | 82.1 | 352 | D72264 | hypothetical prote |
| 11 | 32 | 82.1 | 367 | AD1786 | cell division prot |
| 12 | 32 | 82.1 | 369 | AF1410 | cell division prot |
| 13 | 32 | 82.1 | 672 | C97152 | conjugative trans |
| 14 | 32 | 82.1 | 1364 | 1 AUFPP2 | phosphoribosylamin |
| 15 | 32 | 82.1 | 1900 | S45530 | probable 1-phospha |
| 16 | 31 | 79.5 | 99 | S44632 | f22b7.3 protein- |
| 17 | 31 | 79.5 | 149 | C70878 | hypothetical prote |
| 18 | 31 | 79.5 | 201 | T28706 | hypothetical prote |
| 19 | 31 | 79.5 | 219 | E75143 | phosphoglycolate p |
| 20 | 31 | 79.5 | 307 | T34973 | 5,10-methylenetet |
| 21 | 31 | 79.5 | 356 | C97265 | mannose-1-phosphat |
| 22 | 31 | 79.5 | 360 | S57777 | cysteine proteinas |
| 23 | 31 | 79.5 | 402 | T13614 | N-acetyltransferas |
| 24 | 31 | 79.5 | 406 | D86895 | membrane protein [|
| 25 | 31 | 79.5 | 432 | P83903 | hypothetical prote |
| 26 | 31 | 79.5 | 454 | C86766 | hypothetical prote |
| 27 | 31 | 79.5 | 468 | S37217 | hexon protein - hu |
| 28 | 31 | 79.5 | 526 | 1 VGN5G | spike glycoprotein |
| 29 | 31 | 79.5 | 533 | H71492 | probable hsp-60 - |

ALIGNMENTS

RESULT 1

T33708

hypothetical protein F58E2.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 08-Sep-2000

C:Accession: T33708

R:Goela, D.; Delehaanty, A.

submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid F58E2.

A:Reference number: Z21390

A:Accession: T33708

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-688 <GO>

A:Cross-references: EMBL:AF100659; PIDN:AAC68967.1; GSPDB:GN00022; CESP:F58E2.4

A:Experimental source: strain Bristol N2; clone F58E2

C:Genetics:

A:Gene: CESP:F58E2.4

A:Map position: 4

A:Introns: 228/3; 309/3; 344/2; 602/3

C:Superfamily: Caenorhabditis elegans hypothetical protein F58E2.3

Query Match

Best Local Similarity 87.2%; Score 34; DB 2; Length 688;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9

Db 393 LIYENVGLS 401

RESULT 2

S67298

deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O6771; protein YOR386W

C:Species: Saccharomyces cerevisiae

C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000

C:Accession: S67298; A23964; A24046

R:Deilus, H.; Hebling, U.; Hofmann, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67261

A:Accession: S67298

A:Molecule type: DNA

A:Residues: 1-565 <DE>

A:Cross-references: EMBL:Z75294; NID:G1420830; PIDN:CAA99718.1; PID:G1420831; MIPS:YOR386

A:Experimental source: strain S288C

R:Yasui, A.; Langeveld, S.A.

Gene 36, 349-355, 1985

A:Title: Homology between the photoreactivation genes of Saccharomyces cerevisiae and Esc

A:Reference number: A23964; MUID:86083177; PMID:3000886

A:Accession: A23964

A:Molecule type: DNA

polyposphate kina
probable ppk prote
hexon protein - hu
multidrug resistan
hemagglutinin - In
conserved hypotet
transcription regu
conserved hypotet
MutT-like protein
MutT/nudix family
hypothetical prote
MutT/nudix family
venom allergen ant
hypothetical prote
ABC transporter, A
cyanamide hydratase

A;Residues: 1-76,'A',78-164,'S',166-168,'T',170-199,'S',201-350,'R',352-364,'E',366-472,
A;Cross-references: EMBL:M11578; NID:g172169; PIDN:AAA34875.1; PID:g172170
R;Sancar, G.B.
Nucleic Acids Res. 13, 8231-8246, 1985
A;Title: Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of the PHR1 pho
A;Reference number: A24046; MUID:85067229; PMID:3906569
A;Accession: A24046
A;Molecule type: DNA
A;Residues: 1-565 <S>
A;Cross-references: EMBL:X03183; NID:g4175; PIDN:CAA26944.1; PID:g4176
C;Genetics:
A;Gene: SGD:PHR1
A;Cross-references: SGD:S0005913; MIPS:YOR386W
A;Map position: 15R
A;Superfamily: decyribodipyrimidine photo-lyase
C;Keywords: carbon-carbon lyase

Query Match 84.6%; Score 33; DB 2; Length 565;
Best Local Similarity 55.6%; Pred. No. 44;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 86 RLYDNVGLY 94
:::|||||::

RESULT 3
JN0805
Protein-tyrosine-phosphatase (EC 3.1.3.48) PTPN11, nonreceptor type 11 [validated] - hum
N;Alternate names: BTP-3; protein-tyrosine-phosphatase SHP-2; PTPD; SH-PTP2; SH
C;Species: Homo sapiens (man)
C;Date: 10-Mar-1994 #sequence revision 19-May-1994 #text change 21-Jun-2002
R;Accession: JN0805; A46210; A47386; A47244; S27398; C44929; S31767
R;Bastien, L.; Ramachandran, C.; Liu, S.; Adam, M.
Biochem. Biophys. Res. Commun. 196, 124-133, 1993
A;Title: Cloning, expression and mutational analysis of SH-PTP2, human protein-tyrosine
A;Reference number: JN0805; MUID:94029983; PMID:8218283
A;Accession: JN0805
A;Molecule type: mRNA
A;Residues: 1-593 <BAS>
A;Cross-references: EMBL:X70766; NID:g292406; PIDN:AAA17022.1; PID:g292407
R;Vogel, W.; Lammers, R.; Huang, J.; Ullrich, A.
Science 259, 1611-1614, 1993
A;Title: Activation of a phosphotyrosine phosphatase by tyrosine phosphorylation.
A;Reference number: A46210; MUID:9306095; PMID:7681217
A;Accession: A46210
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-593 <VOG>
A;Cross-references: EMBL:X70766; NID:g35783; PIDN:CAA50045.1; PID:g35784
A;Experimental source: SK-BR-3 mammary carcinoma cells
A;Note: sequence extracted from NCBI backbone (NCBIP:127775)
R;Ahmad, S.; Barville, D.; Zhao, Z.; Fischer, E.H.; Shen, S.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 2197-2201, 1993
A;Title: A widely expressed human protein-tyrosine phosphatase containing src homology 2
A;Reference number: A47386; MUID:93211929; PMID:7681589
A;Accession: A47386
A;Molecule type: mRNA
A;Residues: 1-593 <AHM>
A;Experimental source: umbilical cord
A;Note: sequence extracted from NCBI backbone (NCBIN:128129, NCBIP:128131)
R;Freeman Jr., R.M.; Plutsky, J.; Neel, B.G.
Proc. Natl. Acad. Sci. U.S.A. 89, 11239-11243, 1992
A;Title: Identification of a human src homology 2-containing protein-tyrosine-phosphatase
A;Reference number: A47244; MUID:93087502; PMID:1280823
A;Accession: A47244
A;Molecule type: mRNA
A;Residues: 1-593 <PRE>
A;Cross-references: GB:L03535; NID:g338081; PIDN:AAA36611.1; PID:g338082
A;Note: sequence extracted from NCBI backbone (NCBIN:119760, NCBIP:119761)
R;Adachi, M.; Sekiya, M.; Miyachi, T.; Matsuno, Y.; Imai, K.; Yachi, A.
FEBS Lett. 314, 335-339, 1992
A;Title: Molecular cloning of a novel protein-tyrosine phosphatase SH-PTP3 with sequence

A;Reference number: S27398; MUID:93106179; PMID:1281790
A;Accession: S27398
A;Molecule type: mRNA
A;Residues: 1-534,'R',536-547,'P',549-593 <AD2>
A;Cross-references: DBJ:DL3940; NID:g220071; PIDN:BA02740.2; PID:g4519425
R;Adachi, M.; Sekiya, M.; Arimura, Y.; Takekawa, M.; Itoh, F.; Hinoda, Y.; Imai, K.; Yachi
Cancer Res. 52, 737-740, 1992
A;Title: Protein-tyrosine phosphatase expression in pre-B cell NALM-6.
A;Reference number: A44929; MUID:92119637; PMID:1370651
A;Accession: C44929
A;Molecule type: mRNA
A;Residues: 1-370-450 <ADA>
A;Cross-references: GB:S78088; NID:g243547; PIDN:AA21148.1; PID:g243548
A;Experimental source: pre-B cell NALM-6
A;Note: sequence extracted from NCBI backbone (NCBIN:78088, NCBIP:78089)
A;Note: the authors did not report the entire codon for residue 92
C;Comment: This ubiquitous enzyme plays a critical role in regulating physiological cell
C;Genetics:
A;Gene: GDB:PTPN11
A;Cross-references: GDB:137093; OMIM:176876
A;Map position: 12q24.1-12q24.1
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phosphatase
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F;6-100/Domain: SH2 homology <SH2A>
F;112-214/Domain: SH2 homology <SH2B>
F;273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
F;459/Active site: Cys (phosphocysteine intermediate) #status predicted
F;465/Binding site: substrate phosphate (Arg) #status predicted

Query Match 84.6%; Score 33; DB 1; Length 593;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 578 RYENVGLM 586
:::|||||::

RESULT 4
JCS167
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - chicken
N;Alternate names: phosphotyrosine phosphatase; PTP1D; PTP2c; SH-PTP2; SyP
C;Species: Gallus gallus (chicken)
C;Date: 21-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 16-Jul-1999
C;Accession: JCS167
R;Park, C.Y.; Lamontagne, K.R.; Tonks, N.K.; Hayman, M.J.
Gene 177, 93-97, 1996
A;Title: Cloning and expression of the chicken protein tyrosine phosphatase SH-PTP2.
A;Reference number: JCS167; MUID:97080506; PMID:8921851
A;Contents: erythroblast
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-593 <PAR>
A;Cross-references: GB:U88620; NID:g1054939; PIDN:AAC50049.1; PID:g1054940
C;Comment: This enzyme plays positive roles in mitogenic signaling and early development.
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phosphatase
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F;6-105,112-193/Domain: SH2 #status predicted <SH2A>
F;6-100/Domain: SH2 homology <SH2B>
F;112-214/Domain: SH2 homology <SH2B>
F;273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
F;559-570/Region: proline-rich
F;559-570/Active site: Cys (phosphocysteine intermediate) #status predicted
F;465/Binding site: substrate phosphate (Arg) #status predicted
F;542,547,580/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 84.6%; Score 33; DB 2; Length 593;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
:::|||||::

Db 578 RYENVGLM 586

RESULT 5
A55651
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - African clawed frog
N:Alternate names: SH-PTP2
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A55651
R:Tang, T.L.; Freeman Jr., R.M.; O'Reilly, A.M.; Neel, B.G.; Sokol, S.Y.
Cell 80, 473-483, 1995
A:Title: The SH2-containing protein-tyrosine phosphatase SH-PTP2 is required upstream of
A:Reference number: A55651; MUID:95163101; PMID:7859288
A:Accession: A55651
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-595 <PAN>
A:Cross-references: GB:U15287; NID:G601781; PIDN:AAAG5731.1; PID:G601782
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phosphatase
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F:6-100/Domain: SH2 homology <SH2A>
F:112-214/Domain: SH2 homology <SH2B>
F:273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
F:459/Active site: Cys (phosphocysteine intermediate) #status predicted
F:465/Binding site: substrate phosphate (Arg) #status predicted

Query Match 84.6%; Score 33; DB 1; Length 595;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9

Db 580 RYENVGLM 588

RESULT 6
A53593
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - rat
N:Alternate names: PTPase L1
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A53593; S29281
R:Mei, L.; Doherty, C.A.; Huganir, R.L.
J. Biol. Chem. 269, 12254-12262, 1994
A:Title: RNA splicing regulates the activity of a SH2 domain-containing protein tyrosine
A:Reference number: A53593; MUID:94216346; PMID:7512964
A:Accession: A53593
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-597 <MEI>
A:Cross-references: GB:U05963; NID:9458332; PIDN:AAAL9133.1; PID:G458333
R:Hiraga, A.; Munakata, H.; Hata, K.; Suzuki, Y.; Tsukiki, S.
Eur. J. Biochem. 209, 195-206, 1992
A:Title: Purification and characterization of a rat liver protein-tyrosine phosphatase w
A:Reference number: S29281; MUID:93011127; PMID:1382983
A:Accession: S29281
A:Molecule type: protein
A:Residues: 24-31;36-54;56-89;100-103,'X',105-108,'X',113-120;132-155;179-198;214-233;24
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phosphatase
C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosi
F:6-100/Domain: SH2 homology <SH2A>
F:112-214/Domain: SH2 homology <SH2B>
F:273-514/Domain: protein-tyrosine-phosphatase homology <PTP>
F:463/Active site: Cys (phosphocysteine intermediate) #status predicted
F:469/Binding site: substrate phosphate (Arg) #status predicted

Query Match 84.6%; Score 33; DB 1; Length 597;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9

:::|||||:::

Db 582 RYENVGLM 590

RESULT 7
T20550
hypothetical protein F07C6.4b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T20550; T23678
R:Steward, C.
Submitted to the EMBL Data Library, February 1996
A:Reference number: Z19290
A:Accession: T20550
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-700 <WIL>
A:Cross-references: EMBL:Z69659; PIDN:CAA93486.1; GSPDB:GN00022; CESP:F07C6.4b
A:Experimental source: clone F07C6
R:Lightning, J.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19780
A:Accession: T23678
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-700 <W12>
A:Cross-references: EMBL:Z81102; PIDN:CAB03204.1; GSPDB:GN00022; CESP:F07C6.4b
A:Experimental source: clone M02B1
C:Genetics:
A:Gene: CESP:F07C6.4b
A:Map position: 4
A:Introns: 21/3; 58/2; 111/1; 159/3; 195/3; 272/1; 328/2; 399/2; 423/3; 546/3; 564/1; 612/1

Query Match 84.6%; Score 33; DB 2; Length 700;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7

Db 398 RYENVG 404

RESULT 8
B69944
hypothetical protein yqac - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: B69944
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallerc
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
t.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yaguchi, K.; Yaguchi, K.; Yaguchi, K.
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69944
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-178 <KUN>
A:Cross-references: GB:Z99117; GB:AL009126; NID:G2634966; PIDN:CAB14578.1; PID:e1183866;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yqac

Query Match 82.1%; Score 32; DB 2; Length 178;

Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 129 SLYDNAGME 137
:|||||:

RESULT 9
H85138
hypothetical protein AT4g12900 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: H85138
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85138
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <STO>
A:Cross-references: GB:NC_001268; NID:g7267992; PIDN:CAB78332.1; GSPDB:GNO0140
C:Genetics:
A:Gene: AT4g12900
A:Map position: 4
C:Superfamily: Arabidopsis thaliana hypothetical protein F7A7.100

Query Match 82.1%; Score 32; DB 2; Length 231;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7
Db 182 PLYENIG 188
:|||||:

RESULT 10
D72264
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 28-Jul-2000
C:Accession: D72264
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:9287316; PMID:10360571
A:Accession: D72264
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <ARN>
A:Cross-references: GB:AE000512; NID:g4981904; PIDN:AAD36419.1; PID:g498190
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1348
C:Superfamily: Thermotoga maritima hypothetical protein TM1348

Query Match 82.1%; Score 32; DB 2; Length 352;
Best Local Similarity 55.6%; Pred. No. 42;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 330 RLYEEIGHM 338
:|||||:

RESULT 11
AD1786
cell division protein FtsW homolog lin2834 [imported] - Listeria innocua (strain Clip112
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AD1786
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1786
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-367 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC98060.1; PID:g16415369; GSPDB:GNO0178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2834
C:Superfamily: rod shape-determining protein

Query Match 82.1%; Score 32; DB 2; Length 367;
Best Local Similarity 44.4%; Pred. No. 44;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 316 NIFENIGMT 324
:|||||:

RESULT 12
AF1410
cell division protein FtsW homolog lmo2687 [imported] - Listeria monocytogenes (strain EC
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1410
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1410
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-369 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC00900.1; PID:g16412187; GSPDB:GNO0177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2687
C:Superfamily: rod shape-determining protein

Query Match 82.1%; Score 32; DB 2; Length 369;
Best Local Similarity 44.4%; Pred. No. 44;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 316 NIFENIGMT 324
:|||||:

RESULT 13
C97152
conjugative transfer gene TrsE homolog, ATPase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: C97152
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koorin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97152

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-617 <R>
A:Cross-references: GB:AE001437; PIDN:AAK80006.1; PID:G15025033; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
A:Genetics:
A:Gene: CAC2047

Query Match 82.1%; Score 32; DB 2; Length 617;
Best Local Similarity 55.6%; Pred. No. 80;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
:|||||:
Db 369 QLYENLGIT 377

RESULT 14
AJFFPP
phosphoribosylamine-glycine ligase (EC 6.3.4.13) - fruit fly (Drosophila pseudoobscura)
N:Alternate names: glycine ribonucleotide synthetase (GARSase); glycine ribonucle
N:Contains: phosphoribosylamine-glycine ligase (EC 6.3.4.13); phosphoribosylformylglycin
C:Species: Drosophila pseudoobscura
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Jun-2002
C:Accession: S01204
R:Henikoff, S.; Eghtedarzadeh, M.K.
Genetics 117, 711-725, 1987
A:Title: Conserved arrangement of nested genes at the Drosophila Gart locus.
A:Reference number: S01204; MUID:88112752; PMID:3123310
A:Accession: S01204
A:Molecule type: DNA
A:Residues: 1-1364 <HENS>
A:Cross-references: EMBL:X06285; NID:G9055; PIDN:CAA29611.1; PID:G295787
A>Note: monofunctional phosphoribosylamine-glycine ligase, prepared by alternative splic
C:Genetics:
A:Gene: Gart
A:Cross-references: FlyBase:FBgn00000053
A:Map position: 4 88
A:Introns: 59/3; 142/2; 359/1; 434/2; 575/1; 927/1
C:Superfamily: Drosophila purine synthesis multifunctional protein; phosphoribosylamine-
myltransferase homology
C:Keywords: alternative splicing; cyclo-ligase; methyltransferase; multifunctional enzym
F:4-430/Domain: phosphoribosylamine-glycine ligase homology <PGL>
F:444-775/Domain: phosphoribosylformylglycinamide cyclo-ligase homology <PFC>
F:794-1124/Domain: phosphoribosylformylglycinamide cyclo-ligase homology <PFCL>
F:1158-1351/Domain: phosphoribosylglycinamide formyltransferase homology <PRGF>

Query Match 82.1%; Score 32; DB 1; Length 1364;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7
:|||||:
Db 514 ELYENIG 520

RESULT 15
S45530
probable 1-phosphatidylinositol 4-kinase (EC 2.7.1.67) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein L2142.4; protein YLR305c
C:Species: Saccharomyces cerevisiae
C>Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 21-Jul-2000
C:Accession: S45530; S51437
R:Yoshida, S.; Ohya, Y.; Goebi, M.; Nakano, A.; Anraku, Y.
J. Biol. Chem. 269, 1166-1172, 1994
A:Title: A novel gene, STT4, encodes a phosphatidylinositol 4-kinase in the PKC1 protein
A:Reference number: S45530; MUID:94117423; PMID:8288577
A:Accession: S45530
A:Molecule type: DNA
A:Residues: 1-1900 <YOS>
A:Cross-references: EMBL:D13717; NID:G454206; PIDN:BAA02870.1; PID:G454207
R:Pauley, A.
submitted to the EMBL Data Library, November 1994

A:Description: The sequence of S. cerevisiae cosmid L2142.
A:Reference number: S51437
A:Accession: S51437
A:Molecule type: DNA
A:Residues: 1-1900 <PAU>
A:Cross-references: EMBL:U17247; NID:G577216; PID:G577220; MIPS:YLR305c
C:Genetics:
A:Gene: SGD:STT4
A:Cross-references: SGD:S0004296; MIPS:YLR305c
A:Map position: 12R
C:Keywords: phosphotransferase; transmembrane protein
F:377-393/Domain: transmembrane #status predicted <TMM>

Query Match 82.1%; Score 32; DB 2; Length 1900;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7
:|||||:
Db 141 VLYENIG 147

Search completed: July 15, 2004, 07:29:23
Job time : 12.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:20:47 ; Search time 8 seconds
(without alignments)
58.579 Million cell updates/sec

Title: SEQMOD

Perfect score: 39

Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | length | ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 33 | 84.6 | 519 | 1 ALGG_PSEPK | Q88nc9 pseudomonas |
| 2 | 33 | 84.6 | 536 | 1 ALGG_PSESM | Q887g3 pseudomonas |
| 3 | 33 | 84.6 | 565 | 1 PHR_YEAST | P05066 saccharomyc |
| 4 | 33 | 84.6 | 593 | 1 PTNE_CHICK | Q90887 gallus gall |
| 5 | 33 | 84.6 | 593 | 1 PTNE_HUMAN | Q06124 homo sapien |
| 6 | 33 | 84.6 | 593 | 1 PTNE_RAT | P41499 rattus norv |
| 7 | 32 | 82.1 | 178 | 1 YQAC_BACSU | P45900 bacillus su |
| 8 | 32 | 82.1 | 1364 | 1 PUR2_DROPS | P16340 d trifuncti |
| 9 | 32 | 82.1 | 1900 | 1 STT4_YEAST | P37297 saccharomyc |
| 10 | 31 | 79.5 | 99 | 1 YLM3_CAEEL | P34406 caenorhabdi |
| 11 | 31 | 79.5 | 307 | 1 METE_STRLI | O54235 streptomyce |
| 12 | 31 | 79.5 | 357 | 1 RLAO_METKA | Q8tx50 methanopyru |
| 13 | 31 | 79.5 | 360 | 1 CYSP_HEMSP | P43156 hemerocalli |
| 14 | 31 | 79.5 | 367 | 1 FPPS_CHICK | P08836 gallus gall |
| 15 | 31 | 79.5 | 468 | 1 HEX_ADE31 | P36855 human adeno |
| 16 | 31 | 79.5 | 526 | 1 VGLG_SIGMA | P12647 sigma virus |
| 17 | 31 | 79.5 | 739 | 1 PPK_MYCLE | Q33127 mycobacteri |
| 18 | 31 | 79.5 | 742 | 1 PPK_MYCTU | P95111 mycobacteri |
| 19 | 31 | 79.5 | 919 | 1 HEX_ADE12 | P19900 human adeno |
| 20 | 31 | 79.5 | 1302 | 1 MDR4_DROME | Q00449 drosophila |
| 21 | 30 | 76.9 | 172 | 1 NUDH_VIECH | Q9ku53 vibrio chol |
| 22 | 30 | 76.9 | 172 | 1 NUDH_VIEBV | Q8der5 vibrio vuln |
| 23 | 30 | 76.9 | 174 | 1 NUDH_NEIMA | Q9j178 neisseria m |
| 24 | 30 | 76.9 | 174 | 1 NUDH_NEIMA | Q9j178 neisseria m |
| 25 | 30 | 76.9 | 174 | 1 NUDH_SHEON | Q8en98 shewanella |
| 26 | 30 | 76.9 | 174 | 1 NUDH_VIEBPA | Q878a4 vibrio para |
| 27 | 30 | 76.9 | 205 | 1 VA5_VESQ | P35786 vesputia squ |
| 28 | 30 | 76.9 | 228 | 1 GLUC_COREF | Q8rg15 corynebacte |
| 29 | 30 | 76.9 | 228 | 1 GLUC_COREL | P48244 corynebacte |
| 30 | 30 | 76.9 | 244 | 1 CYAH_MYRVE | P22143 myrothecium |
| 31 | 30 | 76.9 | 291 | 1 ENGC_STAAC | Q9kx08 staphylococ |
| 32 | 30 | 76.9 | 291 | 1 ENGC_STAAM | Q99up7 staphylococ |
| 33 | 30 | 76.9 | 342 | 1 ENGC_HAEDU | Q7vml1 haemophilus |

ALIGNMENTS

| | | | | | | | | | |
|----------|--|-----------|------|--|-----|--|--|--|--|
| RESULT 1 | | | | | | | | | |
| ID | ALGG_PSEPK | STANDARD; | PRT; | 519 | AA. | | | | |
| AC | Q88NC9; | | | | | | | | |
| DT | 10-OCT-2003 (Rel. 42, Created) | | | | | | | | |
| DT | 10-OCT-2003 (Rel. 42, Last sequence update) | | | | | | | | |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update) | | | | | | | | |
| DE | Poly(beta-D-mannuronate) C5 epimerase precursor (EC 5.1.3.-). | | | | | | | | |
| GN | ALGG OR PPI283. | | | | | | | | |
| OS | Pseudomonas putida (strain KT2440). | | | | | | | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; | | | | | | | | |
| OC | Pseudomonadaceae; Pseudomonas. | | | | | | | | |
| OX | NCBI_TaxID=160488; | | | | | | | | |
| RN | [1] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RX | MEDLINE=22423060; PubMed=12534463; | | | | | | | | |
| RA | Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H., | | | | | | | | |
| RA | Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M., | | | | | | | | |
| RA | Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J., | | | | | | | | |
| RA | Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I., | | | | | | | | |
| RA | Chris Lee P., Holtzapfel E., Scanlan D., Tran K., Moazzez A., | | | | | | | | |
| RA | Utterback T., Rizzo M., Lee K., Kosack D., Moesti D., Wedder H., | | | | | | | | |
| RA | Lauber J., Stjepandic D., Hohnes J., Straetz M., Helm S., | | | | | | | | |
| RA | Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B., | | | | | | | | |
| RA | Fraser C.M.; | | | | | | | | |
| RT | "Complete genome sequence and comparative analysis of the | | | | | | | | |
| RT | metabolically versatile Pseudomonas putida KT2440."; | | | | | | | | |
| RL | Environ. Microbiol. 4:799-808(2002). | | | | | | | | |
| CC | -!- FUNCTION: Bifunctional protein that converts poly(beta-D- | | | | | | | | |
| CC | mannuronate) to alpha-L-gulonate and that is also part of a | | | | | | | | |
| CC | periplasmic protein complex that serves as a scaffold that leads | | | | | | | | |
| CC | the newly formed alginate polymer through the periplasmic space to | | | | | | | | |
| CC | the outer membrane secretin alge (By similarity). | | | | | | | | |
| CC | -!- PATHWAY: Alginate biosynthesis. | | | | | | | | |
| CC | -!- SUBCELLULAR LOCATION: Periplasmic (Probable). | | | | | | | | |
| CC | -!- SIMILARITY: Belongs to the D-mannuronate C5-epimerase family. | | | | | | | | |
| CC | -!- SIMILARITY: Contains 6 PBH1 repeats. | | | | | | | | |
| CC | ----- | | | | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | | | | | |
| CC | ----- | | | | | | | | |
| DR | EMBL; AE016778; AAN66907.1; ALT_INIT. | | | | | | | | |
| DR | TIGR; PPI283; - | | | | | | | | |
| DR | InterPro; IPR006633; CASH. | | | | | | | | |
| DR | InterPro; IPR006626; PBH1. | | | | | | | | |
| DR | SMART; SM00722; CASH; 1. | | | | | | | | |
| DR | SMART; SM00710; PBH1; 6. | | | | | | | | |
| KW | Alginate biosynthesis; Isomerase; Periplasmic; Repeat; Signal; | | | | | | | | |
| KW | Complete proteome. | | | | | | | | |
| FT | SIGNAL | 1 | 25 | POTENTIAL. | | | | | |
| FT | CHAIN | 26 | 519 | POLY(BETA-D-MANNURONATE) C5 EPIMERASE. | | | | | |

P50525 schizosacch
P36850 human adeno
Q92h60 rickettsia
P34219 saccharomyc
P03451 influenza a
O60879 homo sapien
O68006 b bacitraci
Q9v1v5 drosophila
Q00812 nostoc comm
P52335 nostoc sp.
P35640 bartonella
Q98f04 rhizobium 1

```

FT REPEAT 219 246 PBH1 1.
FT REPEAT 281 303 PBH1 2.
FT REPEAT 305 328 PBH1 3.
FT REPEAT 330 352 PBH1 4.
FT REPEAT 354 376 PBH1 5.
FT REPEAT 377 399 PBH1 6.
SQ SEQUENCE 519 AA; 57936 MW; 804D0C87D39EDCC CRC64;

Query Match 84.6%; Score 33; DB 1; Length 519;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 404 XLYENVAMA 412

RESULT 2
ALGG PSESMS
ID ALGG PSESMS STANDARD; PRT; 536 AA.
AC Q887Q3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-VAR-2004 (Rel. 43, Last annotation update)
DE Poly(beta-D-mannuronate) C5 epimerase precursor (EC 5.1.3.-).
GN ALGG OR PSPT01238.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwin M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Bean M.J., Haft D.H.,
RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA Van Aken S.E., Feldblum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collmer A.;
RA "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC -!- FUNCTION: Bifunctional protein that converts poly(beta-D-
CC mannuronate) to alpha-L-gulonate and that is also part of a
CC periplasmic protein complex that serves as a scaffold that leads
CC the newly formed alginate polymer through the periplasmic space to
CC the outer membrane secretin alge (By similarity).
CC -!- PATHWAY: Alginate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC -!- SIMILARITY: Belongs to the D-mannuronate C5-epimerase family.
CC -!- SIMILARITY: Contains 5 Pbh1 repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF016860; AAO54763.1; -
CC TIGR; PSPT01238; -
CC SMART; SM00722; CASH; 1.
CC SMART; SM00710; PBH1; 5.
CC Complete proteome.
CC SIGNAL 1 36 POTENTIAL.
CC CHAIN 37 536 POLY(BETA-D-MANNURONATE) C5 EPIMERASE.
CC REPEAT 298 320 PBH1 1.

RESULT 3
PHR YEAST
ID PHR YEAST STANDARD; PRT; 565 AA.
AC P05066;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deoxyribodipyrimidine photolyase, mitochondrial precursor
DE (EC 4.1.99.3) (DNA photolyase) (Photoreactivating enzyme).
GN PHR1 OR YOR186W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067229; PubMed=3906569;
RA Sancar G.B.;
RT "Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of
RL the PHR1 photolyase to E. coli photolyase.";
RL Nucleic Acids Res. 13:8231-8246(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86083177; PubMed=3000886;
RA Yasui A., Langeveld S.A.;
RT "Homology between the photoreactivation genes of Saccharomyces
RL cerevisiae and Escherichia coli.";
RL Gene 36:349-355(1985).
RN [3]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U., Hofmann B.;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBSJ databases.
RN [4]
RP REVIEW.
RA Sancar G.B., Sancar A.;
RT "Structure and function of DNA photolyases.";
RL Trends Biochem. Sci. 12:259-261(1987).
CC -!- FUNCTION: This enzyme catalyzes the light-dependent monomerization
CC (300-600 nm) of cyclobutyl pyrimidine dimers (in cis-syn
CC configuration), which are formed between adjacent bases on the
CC same DNA strand, upon exposure to ultraviolet radiation.
CC -!- CATALYTIC ACTIVITY: Cyclobutadipyrimidine (in DNA) = 2 pyrimidine
CC residues (in DNA).
CC -!- COFACTOR: Contains 2 chromophores: a reduced flavin (FADH2) and a
CC 5,10-methenyltetrahydrofolate. Both chromophores are bound by non-
CC covalent interactions.
CC -!- SUBCELLULAR LOCATION: Nuclear and mitochondrial.
CC -!- MISCELLANEOUS: This protein belongs to the "short wavelength-type
CC photolyases" with an absorption maximum at about 380 nm.
CC -!- MISCELLANEOUS: There are only 150-300 molecules of photolyase per
CC yeast cell.
CC -!- SIMILARITY: Belongs to the DNA photolyase class-1 family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
-----CC
```

| | | |
|-------|--|---------------------------------|
| | EMLB; X03183; CAA26944.1; | -- |
| DR DR | EMLB; M11578; AAA34875.1; | -- |
| DR DR | EMLB; ZV5294; CAA93718.1; | -- |
| DR DR | PIR; S67298; S67298. | |
| DR DR | HSSP; P00914; IDNP. | |
| DR DR | GermOnline; I43974. | -- |
| DR DR | SGD; S0005913; PHR1. | |
| DR DR | InterPro; IPRO02081; DNA_photolyase_1. | |
| DR DR | InterPro; IPRO06050; DNA_photolyase_N. | |
| DR DR | InterPro; IPRO05101; FAD_Binding_7. | N |
| DR DR | InterPro; IPRO06051; FAD_Binding_N. | |
| DR DR | Pfam; PF00875; DNA_photolyase_1. | |
| DR DR | Pfam; PF03441; FAD_binding_7; 1. | |
| DR DR | PRINTS; PR00147; DNAPHOTOLYASE. | |
| DR DR | ProDom; PD004390; FAD_binding_N_1. | |
| DR DR | FPROSITE; PS00394; DNA_PHOTOLYASES_1_1; 1. | |
| DR DR | FPROSITE; PS00691; DNA_PHOTOLYASES_1_2; 1. | |
| DR KW | Lysine; Chromophore; Flavoprotein; FAD; 1. | DNA repair; DNA-binding; |
| DR KW | Nuclear protein; Mitochondrion; Transit peptide. | |
| FT FT | TRANSIT ? MITOCHONDRION | |
| FT FT | CHAIN ? | DEOXYRIBODIPYRIDINE PHOTOLYASE. |
| FT FT | DNA_BIND 421 440 | (POTENTIAL). |
| FT FT | CONFLICT 77 77 | V -> A (IN REF. 2). |
| FT FT | CONFLICT 165 165 | T -> S (IN REF. 2). |
| FT FT | CONFLICT 169 169 | S -> T (IN REF. 2). |
| FT FT | CONFLICT 200 200 | D -> S (IN REF. 2). |
| FT FT | CONFLICT 351 351 | S -> R (IN REF. 2). |
| FT FT | CONFLICT 365 365 | G -> E (IN REF. 2). |
| FT FT | CONFLICT 473 473 | E -> K (IN REF. 2). |
| FO SQ | SEQUENCE 565 AA; 662274 MW; | CD4FC3DA6128B97C CRC64; |

| | | | | |
|-------------------------|--------|---------------|-----------|-------------|
| Query Match | 84.6% | Score 33; | DB 1; | Length 565; |
| Best Local Similarity | 55.6%; | Pred. No. 30; | | |
| Matches 5; Conservative | 4; | Mismatches | 0; Indels | 0; Caps |
| | | | | 0; |

| | | | |
|----|----|-----------|----|
| QY | 1 | XYENVMGX | 9 |
| | | : : : : : | |
| Db | 86 | FLYDNVGLY | 94 |

RESULT 4
PTNB_CHICK
ID_PTNB CHICK
STANDARD: PRT: 593 AA.

AC Q90687;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase, non-receptor type 11 (EC 3.1.3.48) (cSH-
DE FTP2).

GN PTPN11 OR SH-PTP2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CX NCBI TaxID=9031;

[1]
RN SEQUENCE FROM N. A.
RP
RN TISSUE=Erythroblast;
RX MEDLINE=9708056; PubMed=8921851;
RC Park C.Y., LaMontagne K.R., Tonks N.K., Hayman M.J.:
RT "Cloning and expression of the chicken protein tyrosine phosphatase
RL SH-PTP2". RT
RL Gene 177:93-97(1996).

-!- FUNCTION: This ptpase activity may directly link growth factor receptors and other signaling proteins through protein-tyrosine phosphorylation. The SH2 regions may interact with other cellular components to modulate its own phosphatase activity against interacting substrates [35] (phosphatase activity against catalytic substrates of erythroid cell proliferation).

-!- CATALYTIC ACTIVITY: protein tyrosine phosphatase + H₂O = protein

Db 578 RVYENVGLM 586

RESULT 7

YQAC_BACSU STANDARD; PRT; 178 AA.

ID P45900;

DT 01-NOV-1995 (Rel. 32, Created)

DT 10-OCT-2003 (Rel. 32, Last sequence update)

DE Hypothetical protein yqac precursor.

GN YQAC OR BSU26370.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RX MEDLINE=95219086; PubMed=7704261;

RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;

RT "Complete nucleotide sequence of a skin element excised by DNA

RT rearrangement during sporulation in *Bacillus subtilis*.";

RL Microbiology 141:323-327(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RX MEDLINE=97124195; PubMed=8969508;

RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,

RA Kobayashi Y.;

RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of

RT the *Bacillus subtilis* genome containing the skin element and many

RT sporulation genes.";

RL Microbiology 142:3103-3111(1996).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertsch M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Broutelle S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,

RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Katamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lepidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara K., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,

RA Sato T., Scallan E., Schleich S., Schroeter R., Scofield F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpatra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,

RA Winters P., Wipat K., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*

RT *subtilis*.";

RL Nature 390:249-256(1997).

RN [4]

RP IDENTIFICATION.

RX MEDLINE=96084975; PubMed=7489895;

RA Medigue C., Moszer I., Viari A., Danchin A.;

RT "Analysis of a *Bacillus subtilis* genome fragment using a co-operative

computer system prototype.";

Gene 165:GC37-GC51(1995).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D32216; BAA06916.1; -

DR EMBL; D84432; BAA12377.1; -

DR EMBL; Z99117; CAB14578.1; -

DR PIR; B69944; B69944.

DR Subtilist; BG11254; yqac.

KW Hypothetical protein; Signal; Complete proteome.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 178 HYPOTHETICAL PROTEIN YQAC.

SQ SEQUENCE 178 AA; 20702 MW; DD2DE09D65CF882E CRC64;

Query Match 82.1%; Score 32; DB 1; Length 178;

Best Local Similarity 55.6%; Pred. No. 14;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9

Db 129 SLYDNAGME 137

RESULT 8

PUR2_DROPS STANDARD; PRT; 1364 AA.

ID PUR2_DROPS

AC P16340;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Trifunctional purine biosynthetic protein adenosine-3 [Includes:

DE Phosphoribosylamine-glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide

DE ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase);

DE Phosphoribosylformylglycinamide cyclo-ligase (EC 6.3.3.1) (AIRS)

DE (Phosphoribosyl-aminimidazole synthetase) (AIR synthase);

DE Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2) (GART) (GAR

DE transformylase) (5'-phosphoribosylglycinamide transformylase)].

AD E3 OR GART.

GN Drosophila pseudoobscura (Fruit fly).

OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7237;

RL [1]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RN STRAIN=EST10;

RX MEDLINE=88112752; PubMed=3123310;

RA Henikoff S., Eghtedarzadeh M.K.;

RT "Conserved arrangement of nested genes at the *Drosophila* Gart locus.";

RL Genetics 117:711-725(1987).

CC -! CATALYTIC ACTIVITY: ATP + 5-phospho-D-ribose + glycine = ADP

CC + phosphate + N(1)-(5-phospho-D-ribose)glycinamide.

CC -! CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + N(1)-(5-phospho-D-

CC ribosyl)glycinamide = tetrahydrofolate + N(2)-formyl-N(1)-(5-

CC phospho-D-ribose)glycinamide.

CC -! CATALYTIC ACTIVITY: ATP + 2-(formamido)-N(1)-(5-phospho-D-

CC ribosyl)acetamide = ADP + phosphate + 5-amino-1-(5-phospho-D-

CC ribosyl)imidazole.

CC -! PATHWAY: De novo purine biosynthesis; second step.

CC -! PATHWAY: De novo purine biosynthesis; third step.

CC -! PATHWAY: De novo purine biosynthesis; fifth step.

CC -! ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=Long;

CC IsoID=P16340-1; Sequence=Displayed;

CC Name=Short;

```
CC IsoId=PI6340-2; Sequence=VSP_005514, VSP_005515;
CC -!- SIMILARITY: In the N-terminal section; belongs to the GARS family.
CC -!- SIMILARITY: In the central section; belongs to the AIR synthase
CC family.
CC -!- SIMILARITY: TO OTHER AIRS AND GART FROM BACTERIA AND EUKARYOTES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X6285; CAA29611.1; -
CC PIR; S01204; AJFFP.
CC HSP; P08179; IGAR.
CC FlyBase; FBgn0012717; Dpse\ade3.
CC InterPro; IPR000728; AIR synth.
CC InterPro; IPR002376; formyl_transf.
CC InterPro; IPR000115; Gars.
CC InterPro; IPR001555; GART AS.
CC InterPro; IPR004733; PurM clligase.
CC InterPro; IPR004607; PurN.
CC Pfam; PF00586; AIRS; 2.
CC Pfam; PF02769; AIRS; 2.
CC Pfam; PF00551; formyl_transf; 1.
CC Pfam; PF01071; GARS; 1.
CC Pfam; PF02842; GARS; 1.
CC Pfam; PF02843; GARS; 1.
CC Pfam; PF02844; GARS; 1.
CC TIGRFAMS; TIGR00877; purD; 1.
CC TIGRFAMS; TIGR00878; purM; 2.
CC TIGRFAMS; TIGR00639; purN; 1.
CC PROSITE; PS00184; GARS; 1.
CC PROSITE; PS00373; GART; 1.
CC KW Multifunctional enzyme; Purine biosynthesis; Ligase; Transferase;
CC Alternative splicing.
CC -----
FT DOMAIN 1 434
FT GARS.
FT DOMAIN 435 1154
FT AIRS.
FT DOMAIN 1155 1364
FT GART.
FT ACT_SITE 1301 1301
FT BY SIMILARITY.
FT VARSPLIC 434 434
FT I -> M (in isoform Short).
FT /FTId=VSP_005514.
FT Missing (in isoform Short).
FT /FTId=VSP_005515.
FT VARSPLIC 435 1364
FT BBD4B5166FF4D301 CRC64;
FT SEQUENCE 1364 AA; 145693 MW; 145693 MW;
Query Match 82.1%; Score 32; DB 1; Length 1364;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 XLVENVG 7
Db 514 ELVENIG 520
RESULT 9
STT4 YEAST STANDARD; PRT; 1900 AA.
AC P37297;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol 4-kinase STT4 (EC 2.7.1.67) (PI4-kinase)
DE (PtdIns-4-kinase).
GN STT4 OR YLR305C OR L2142.4.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=S288c;
RX MEDLINE=94117423; PubMed=8288577;
RA Yoshida S., Goebel M., Ohya Y., Nakano A., Anraku Y.;
RT "A novel gene, STT4, encodes a phosphatidylinositol 4-kinase in the
RT PKC1 protein kinase pathway of Saccharomyces cerevisiae.";
RL J. Biol. Chem. 269:1166-1172(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=973113267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoerge W.,
RA Benes V., Bruckner M., Delius H., Dubois E., Duesterhoef A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwart S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wiedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
CC -!- FUNCTION: Acts on phosphatidylinositol (PI) in the first
CC committed step in the production of the second messenger
CC inositol-1,4,5,-trisphosphate. STT4 functions in PKC1 protein
CC kinase pathway.
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-ID-myo-inositol = ADP +
CC 1-phosphatidyl-ID-myo-inositol 4-phosphate.
CC -!- SIMILARITY: Belongs to the PI3/P14-kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D13717; BAA02870.1; -
CC EMBL; U17247; AAB67358.1; -
CC EMBL; U17243; AAB67354.1; -
CC FIR; S45530; S45530.
CC Germonline; 142368; -.
CC SGD; S0004296; STT4.
CC GO; GO:0005886; C:plasma membrane; IDA.
CC GO; GO:0004430; F:1-phosphatidylinositol 4-kinase activity; IMP.
CC GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; IMP.
CC GO; GO:0000165; P:MAPKK cascade; IDA.
CC GO; GO:0006646; P:phosphatidyethanolamine biosynthesis; IMP.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR00403; PI3_P14_kinase.
CC InterPro; IPR001263; PI3Ka.
CC Pfam; PF00454; PI3_P14_kinase; 1.
CC Pfam; PF00613; PI3Ka; 1.
CC SMART; SM00145; PI3Ka; 1.
CC SMART; SM00146; PI3Kc; 1.
CC PROSITE; PS00915; PI3_4_KINASE_1; 1.
CC PROSITE; PS00916; PI3_4_KINASE_2; 1.
CC PROSITE; PS0290; PI3_4_KINASE_3; 1.
CC Transferase; Kinase.
FT DOMAIN 1643 1882 PI3K/PI4K.
FT SEQUENCE 1900 AA; 214605 MW; F210BAP987BA276A CRC64;
Query Match 82.1%; Score 32; DB 1; Length 1900;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 XLVENVG 7
Db 141 VLYENIG 147
```



```

CC EMBL; AE010373; AAM02039.1; ALT_INIT.
DR HAMAP; MF 00280; -; 1.
DR InterPro; IPR001790; Ribosomal L10.
DR Pfam; PF00466; Ribosomal L10; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 357 AA; 39250 MW; 470294320ADBBESC CRC64;

Query Match
Best Local Similarity 79.5%; Score 31; DB 1; Length 357;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 YENVGXK 9
DB 35 YENVGIV 41

RESULT 13
CISP_HEMSP
ID CISP_HEMSP STANDARD; PRT; 360 AA.
AC P43156;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thiol protease SEN102 precursor (EC 3.4.22.-).
GN SEN102.
OS Hemerocallis sp. (Daylily).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
OC Hemerocallidaceae; Hemerocallis.
OX NCBI_TaxID=29711;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Cradle Song; TISSUE=Petal;
RX MEDLINE=95359413; PubMed=7632925;
RA Valpuesta V., Lange N., Guerrero C., Reid M.;
RT "Up-regulation of a cysteine protease accompanies the ethylene-
RT insensitive senescence of daylily (Hemerocallis) flowers.";
RL Plant Mol. Biol. 28:575-582(1995).
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).
CC -!- SIMILARITY: Belongs to peptidase family C1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X74406; CAA52425.1; -.
DR PIR; S57777; S57777.
DR HSP; P07711; ICUL.
DR MEROPS; C01.010; -.
DR InterPro; IPR000886; ER target S.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPA1N.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SMO0645; Pept_C1; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00139; THIOI PROTEASE CYS; 1.
DR PROSITE; PS00639; THIOI PROTEASE HIS; 1.
DR PROSITE; PS00640; THIOI PROTEASE ASN; 1.
KW Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal;
KW Endoplasmic reticulum.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 133 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 134 360 THIOI PROTEASE SEN102.
FT ACT_SITE 154 154 BY SIMILARITY.
FT ACT_SITE 289 289 BY SIMILARITY.
FT ACT_SITE 310 310 BY SIMILARITY.

```

```

FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SITE 357 360 PREVENT SELECTION FROM ER (POTENTIAL).
SQ SEQUENCE 360 AA; 39914 MW; 808A3D52D2A2C63 CRC64;

Query Match
Best Local Similarity 79.5%; Score 31; DB 1; Length 360;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7
DB 121 FMYENVG 127

RESULT 14
PPPS_CHICK
ID PPPS_CHICK STANDARD; PRT; 367 AA.
AC P08836;
DT 01-NOV-1988 (Rel. 09, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Farnesyl pyrophosphate synthetase (FPP synthetase) (FPP) (Farnesyl
DE diphosphate synthetase) [Includes: Dimethylallyltransferase
DE (EC 2.5.1.1); Geranyltransferase (EC 2.5.1.10)].
GN FDP5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE OF 187-216.
RC TISSUE=Liver;
RX MEDLINE=82000466; PubMed=7272273;
RA Brems D.N., Bruenger E., Rillings H.C.;
RT "Isolation and characterization of a photoaffinity-labeled peptide
RT from the catalytic site of prenyltransferase.";
RL Biochemistry 20:3711-3718(1981).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 20-367.
RC TISSUE=Liver;
RX MEDLINE=94368786; PubMed=8086404;
RA Tarsis L.C., Yan M., Poulter C.D., Sacchettini J.C.;
RT "Crystal structure of recombinant farnesyl diphosphate synthase at
RT 2.6-A resolution.";
RL Biochemistry 33:10871-10877(1994).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-367.
RC TISSUE=Liver;
RX MEDLINE=97140274; PubMed=8986756;
RA Tarsis L.C., Proteau P.J., Kellogg B.A., Sacchettini J.C.,
RA Poulter C.D.;
RT "Regulation of product chain length by isoprenyl diphosphate
RT synthases.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:15018-15023(1996).
CC -!- FUNCTION: Catalyzes the sequential condensations of isopentenyl
CC pyrophosphate with the allylic pyrophosphates, dimethylallyl
CC pyrophosphate, and then with the resultant geranylpyrophosphate to
CC the ultimate product farnesyl pyrophosphate.
CC -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC diphosphate = diphosphate + geranyl diphosphate.
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -!- PATHWAY: Isoprene biosynthesis, cholesterol biosynthesis.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the FPP/GGPP synthetase family.
DR PDB; 1FPS; 10-JUL-95.
DR PDB; 1UBV; 12-MAR-97.
DR PDB; 1UEW; 12-MAR-97.
DR PDB; 1UBX; 12-MAR-97.
DR PDB; 1UBI; 12-MAR-97.
DR InterPro; IPR000092; Polyprenyl_synth.
DR InterPro; IPR008949; Terpenoid_synth.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:25:27 ; Search time 33 Seconds
(without alignments)
86.050 Million cell updates/sec

Title: SEQIMOD

Perfect score: 39

Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 34 | 87.2 | 342 | 13 Q91B95 | Q91B95 potamotrygo |
| 2 | 34 | 87.2 | 430 | 16 Q91R14 | Q91R14 streptomyc |
| 3 | 34 | 87.2 | 688 | 5 Q91Z8 | Q91Z8 caenorhabdi |
| 4 | 33 | 84.6 | 56 | 16 Q89RP9 | Q89RP9 bradyrhizob |
| 5 | 33 | 84.6 | 594 | 13 Q72W17 | Q72W17 brachydanio |
| 6 | 33 | 84.6 | 595 | 13 Q92124 | Q92124 xenopus lae |
| 7 | 33 | 84.6 | 597 | 11 Q64509 | Q64509 mus musculu |
| 8 | 33 | 84.6 | 613 | 16 Q89P36 | Q89P36 bradyrhizob |
| 9 | 33 | 84.6 | 700 | 5 P90329 | P90329 caenorhabdi |
| 10 | 33 | 84.6 | 1353 | 5 Q9VM53 | Q9VM53 drosophila |
| 11 | 32 | 82.1 | 133 | 16 Q88GB4 | Q88GB4 pseudomonas |
| 12 | 32 | 82.1 | 231 | 10 Q9SV79 | Q9SV79 arabidopsis |
| 13 | 32 | 82.1 | 283 | 8 Q8SML6 | Q8SML6 dunaliella |
| 14 | 32 | 82.1 | 352 | 16 Q9X169 | Q9X169 thermotoga |
| 15 | 32 | 82.1 | 365 | 16 Q83BB7 | Q83BB7 thermanser |
| 16 | 32 | 82.1 | 367 | 16 Q927F5 | Q927F5 listeria in |

| | | | | | |
|----|----|------|------|-----------|---------------------|
| 17 | 32 | 82.1 | 369 | 16 Q8Y3Z2 | Q8Y3Z2 listeria mo |
| 18 | 32 | 82.1 | 403 | 16 Q88VH6 | Q88VH6 lactobacilli |
| 19 | 32 | 82.1 | 453 | 16 Q8DVF8 | Q8DVF8 streptococc |
| 20 | 32 | 82.1 | 610 | 16 Q88EC4 | Q88EC4 pseudomonas |
| 21 | 32 | 82.1 | 617 | 16 Q97HG4 | Q97HG4 clostridium |
| 22 | 32 | 82.1 | 783 | 16 Q839N9 | Q839N9 enterococcu |
| 23 | 32 | 82.1 | 1307 | 5 Q8MT77 | Q8MT77 drosophila |
| 24 | 32 | 82.1 | 1817 | 13 Q7SZF6 | Q7SZF6 xenopus lae |
| 25 | 31 | 79.5 | 102 | 16 Q82VV7 | Q82VV7 nitrosomona |
| 26 | 31 | 79.5 | 135 | 11 Q7TN07 | Q7TN07 mus musculu |
| 27 | 31 | 79.5 | 149 | 16 Q33283 | Q33283 mycobacteri |
| 28 | 31 | 79.5 | 149 | 16 Q7TY00 | Q7TY00 mycobacteri |
| 29 | 31 | 79.5 | 165 | 16 Q7VH04 | Q7VH04 helicobacte |
| 30 | 31 | 79.5 | 166 | 5 Q01517 | Q01517 caenorhabdi |
| 31 | 31 | 79.5 | 169 | 2 Q9XDA0 | Q9XDA0 clostridium |
| 32 | 31 | 79.5 | 183 | 11 Q8CES0 | Q8CES0 mus musculu |
| 33 | 31 | 79.5 | 219 | 17 Q9V1X0 | Q9V1X0 pyrococcus |
| 34 | 31 | 79.5 | 221 | 10 Q8W437 | Q8W437 vigna radia |
| 35 | 31 | 79.5 | 282 | 16 Q83B64 | Q83B64 coxiella bu |
| 36 | 31 | 79.5 | 307 | 16 Q9SZV3 | Q9SZV3 streptomyc |
| 37 | 31 | 79.5 | 307 | 16 Q82AF8 | Q82AF8 streptomyc |
| 38 | 31 | 79.5 | 321 | 16 Q8CQ89 | Q8CQ89 staphylococ |
| 39 | 31 | 79.5 | 356 | 16 Q97EV7 | Q97EV7 clostridium |
| 40 | 31 | 79.5 | 377 | 5 Q95RC0 | Q95RC0 drosophila |
| 41 | 31 | 79.5 | 383 | 5 Q9W5A3 | Q9W5A3 drosophila |
| 42 | 31 | 79.5 | 401 | 16 Q88VY0 | Q88VY0 lactobacilli |
| 43 | 31 | 79.5 | 402 | 5 Q46309 | Q46309 drosophila |
| 44 | 31 | 79.5 | 406 | 16 Q9CDQ4 | Q9CDQ4 lactococcus |
| 45 | 31 | 79.5 | 429 | 16 Q834E7 | Q834E7 enterococcu |

ALIGNMENTS

RESULT 1

| | | | |
|----------|---|------|---------|
| Q91B95 | PRELIMINARY; | PRT; | 342 AA. |
| ID | Q91B95; | | |
| AC | Q91B95; | | |
| DT | 01-OCT-2000 (TRENBLrel. 15, Created) | | |
| DT | 01-OCT-2000 (TRENBLrel. 15, Last sequence update) | | |
| DT | 01-JUN-2003 (TRENBLrel. 24, Last annotation update) | | |
| DE | RYTPN6b protein (Fragment). | | |
| GN | RYTPN6B. | | |
| OS | Potamotrygon motoro (South American freshwater stingray). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; | | |
| OC | Elasmobranchii; Squala; Hymnosqualea; Pristiogastera; Batoidae; | | |
| OC | Myliobatiformes; Myliobatoidei; Potamotrygonidae; Potamotrygon. | | |
| OX | NCBI_TaxID=86373; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=20219325; PubMed=10754074; | | |
| RA | Ono-Koyanagi K., Suga H., Katoh K., Miyata T.; | | |
| RT | "Protein tyrosine phosphatases from amphioxus, hagfish, and ray; | | |
| RT | divergence of tissue-specific isoform genes in the early evolution of | | |
| RT | vertebrates." | | |
| RL | J. Mol. Evol. 50:302-311(2000). | | |
| DR | EMBL; AB033591; BAA95198.1; -. | | |
| DR | HSP; Q06124; 2SHP. | | |
| DR | GO; GO:0016787; F:hydrolase activity; IEA. | | |
| DR | GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA. | | |
| DR | GO; GO:0006470; P:protein amino acid dephosphorylation; IEA. | | |
| DR | InterPro; IPR000387; Tyr_PTPase. | | |
| DR | InterPro; IPR000242; Tyr_PP. | | |
| DR | Pfam; PF00102; Y_phosphatase; 1. | | |
| DR | PRINTS; PR00700; PRTYPHPTASE. | | |
| DR | SMART; SM00194; PTPC; 1. | | |
| DR | PROSITE; PS00383; TYR_PHOSPHATASE_1; 1. | | |
| DR | PROSITE; PS50056; TYR_PHOSPHATASE_2; 1. | | |
| DR | PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1. | | |
| FT | NON TER | | |
| KW | 1 | | |
| SEQUENCE | 342 AA; 39532 MW; FCAEEA69442A4677 CRC64; | | |

```

Query Match      87.2%; Score 34; DB 13; Length 342;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 327 RIYENVGLM 335
:::|||||:::

RESULT 2
Q9RIV4 PRELIMINARY; PRT; 430 AA.
AC Q9RIV4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative solute-binding protein.
GN SCO0952 OR SCW11.07C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., James K.D., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Hopwood D.A.;
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
RL EMBL; AL939107; CAB61918.1;
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
KW Complete proteome.
SQ SEQUENCE 430 AA; 46312 MW; 151F92EBF5B9C754 CRC64;

Query Match      87.2%; Score 34; DB 16; Length 430;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 288 NUYENIGIT 296
:::|||||:::

```

```

RESULT 3
Q9IYZ8 PRELIMINARY; PRT; 688 AA.
ID Q9IYZ8;
AC Q9IYZ8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F58E2.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Goela D., Delehaanty A.;
RT "The sequence of C. elegans cosmid F58E2.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL; AF100659; AAC68967.1;
DR PIR; T33708; T33708.
DR WormPep; F58E2.4; CEI7132.
DR InterPro; IPR002900; DUF38.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01827; FTH; 2.
KW Hypothetical protein.
SQ SEQUENCE 688 AA; 79592 MW; 338530655E757124 CRC64;

Query Match      87.2%; Score 34; DB 5; Length 688;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 393 LIYENVGLS 401
:::|||||:::

RESULT 4
Q89RP9 PRELIMINARY; PRT; 56 AA.
ID Q89RP9;
AC Q89RP9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BSL2713 protein.
GN BSL2713.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsurioaka H., Wada T., Yamada M.,

```

01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
Phenothiazyl-protein phosphatase (EC 3.1.3.48).
Xenopus laevis (African clawed frog).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
NCBI_TaxID=8355;
[1]
SEQUENCE FROM N.A.
TISSUE=Oocyte;
MEDLINE=95163101; PubMed=7859288;
Tang T.L., Freeman R.M. Jr., O'Reilly A.M., Neel B.G., Sokol S.Y.;
"The SH2-containing protein-tyrosine phosphatase SH-PTP2 is required
upstream of MAP kinase for early Xenopus development.";
Cell 80:473-483(1995).
EMBL; U15287; AAA65731.1; -.
PIR; A55651; A55651.
HSSP; P35235; IAYA.
GO; GO:0016787; F:hydrolase activity; IEA.
GO; GO:0004727; F:phosphorylated protein tyrosine phosphatase act. . .; IEA.
GO; GO:0007242; P:intracellular signaling cascade; IEA.
GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
InterPro; IPR000387; TYR_phosphatase.
InterPro; IPR00242; Tyr_Pp.
Pfam; PF00102; Y_phosphatase; 1.
PRINTS; PR00700; PTYPHPTASE.
PRINTS; PR00401; SH2DOMAIN.
ProDom; PD000093; SH2; 2.
SMART; SM00194; PTPc; 1.
SMART; SM00252; SH2; 2.
PROSITE; PS50001; SH2; 2.
PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
Hydrolase.
SEQUENCE 595 AA; 68249 MW; 4F39FCF1E8F8D726 CRC64;
Query Match 84.6%; Score 33; DB 13; Length 595;
Best Local Similarity 55.6%; Pred. No. 3.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
:
1 XLYENVGMX 9
:
580 RVYENVGLL 588
PRELIMINARY; PRT; 597 AA.
RESULT 7
ID Q64509
Q64509 PRELIMINARY; PRT; 597 AA.
Q64509
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Protein tyrosine phosphatase (EC 3.1.3.48).
PTPN11.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=Balb/c; TISSUE=Brain;
MEDLINE=96411777; PubMed=8810330;
Ohnishi H., Kubota M., Ohtake A., Sato K., Sano S.;
"Activation of protein-tyrosine phosphatase SH-PTP2 by a tyrosine-
based activation motif of a novel brain molecule.";
J. Biol. Chem. 271:25569-25574(1996).
EMBL; D84372; BAA12328.1; -.
HSSP; P35235; IAYA.

```

DR MGD; MGI:99511; Ptpn11.
DR GO; GO:0005515; P:protein binding; IPI.
DR GO; GO:0007409; P:axoogenesis; IMP.
DR GO; GO:0048011; P:NGF receptor signaling pathway; IMP.
DR InterPro; IPR000380; SH2.
DR InterPro; IPR000387; Tyr phosphatase.
DR InterPro; IPR000422; Tyr_PP.
DR Pfam; PF00107; SH2; 2.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPTASE.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 2.
DR SMART; SM00194; ETPC; 1.
DR SMART; SM00252; SH2; 2.
DR PROSITE; PS00001; SH2; 2.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00566; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00555; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 597 AA; 68460 MW; C742BED37E39EA23 CRC64;

Query Match 84.6%; Score 33; DB 11; Length 597;
Best Local Similarity 55.6%; Pred.No. 3.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9
Db 582 RVYENVGLM 590

RESULT 8
Q89P36 PRELIMINARY; PRT; 613 AA.
AC Q89P36;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B113647 protein.
GN B113647.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL DNA Res. 9:189-197(2002).
DR EMBL; AF005948; BAC48912.1; -.
DR GO; GO:0009058; P:biosynthesis; IPA.
DR InterPro; IPR001296; Glyco_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Complete proteome.
SQ SEQUENCE 613 AA; 68932 MW; 53226C6AD8B83AE1 CRC64;

Query Match 84.6%; Score 33; DB 16; Length 613;
Best Local Similarity 85.7%; Pred.No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7
Db 317 SLYENVG 323

RESULT 9
P90929 PRELIMINARY; PRT; 700 AA.
ID P90929
AC P90929;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F07C6 4b protein.
GN F07C6.4 OR F07C6.4B
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lightning J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81102; CAB03204.1; -.
DR EMBL; Z89659; CAB03204.1; JOINED.
DR EMBL; Z89659; CAB03204.1; -.
DR EMBL; Z81102; CAB03204.1; JOINED.
DR PIR; T20550; T20550.
DR WormBep; F07C6.4b; CE18569.
KW Hypothetical protein.
SQ SEQUENCE 700 AA; 77598 MW; 293869E242E3C6DA CRC64;

Query Match 84.6%; Score 33; DB 5; Length 700;
Best Local Similarity 85.7%; Pred.No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVG 7
Db 398 PLYENVG 404

RESULT 10
Q9VM53 PRELIMINARY; PRT; 1353 AA.
ID Q9VM53;
AC Q9VM53;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG31628 protein.
GN CG31628 OR CG8761.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

```

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Ketchum K.A.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Tobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "the genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Chape M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ignotz M., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Murphy C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Parag V., Park S., Patel S., Pfeiffer B., Nunoo J.,
 RA Prounanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of *Drosophila melanogaster* genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaninker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Asburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
 RT "Annotation of *Drosophila melanogaster* genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003615; AAF52474.2; -;
 DR HSSP; P08179; 1GAR;
 DR FlyBase; FBgn0000053; ade3;
 DR GO; GO:0005737; C:cytoplasm; IEA;
 DR GO; GO:0003824; C:catalytic activity; IEA;
 DR GO; GO:0004637; F:phosphoribosylamine-glycine ligase activity; IEA;
 DR GO; GO:0004641; F:phosphoribosylformylglycinamide cyclo-lig. . . ; IEA;
 DR GO; GO:0004644; F:phosphoribosylglycinamide formyltransferase. . . ; IEA;
 DR GO; GO:0006189; P:de novo IMP biosynthesis; IEA;
 DR GO; GO:0009058; P:biosynthesis; IEA.

DR GO; GO:0009113; P:purine base biosynthesis; IEA.
 DR InterPro; IPR000728; AIR_synth.
 DR InterPro; IPR002376; formyl_transf.
 DR InterPro; IPR000115; Gars.
 DR InterPro; IPR001555; GART_AS.
 DR InterPro; IPR004733; PurM_cligase.
 DR InterPro; IPR004607; PurN.
 DR Pfam; PF00586; AIRS_C; 2.
 DR Pfam; PF02769; AIRS_C; 2.
 DR Pfam; PF00551; formyl_transf; 1.
 DR Pfam; PF01071; GARS; 1.
 DR Pfam; PF02842; GARS_B; 1.
 DR Pfam; PF02843; GARS_C; 1.
 DR Pfam; PF02844; GARS_N; 1.
 DR TIGRFAMs; TIGR00877; purD; 1.
 DR TIGRFAMs; TIGR00878; purM; 2.
 DR TIGRFAMs; TIGR00639; PurN; 1.
 DR PROSITE; PS00184; GARS; 1.
 DR PROSITE; PS00373; GART; 1.
 SQ SEQUENCE 1353 AA; 144525 MW; 3F193005CFID7ACB CRC64;
 Query Match 84.6%; Score 33; DB 5; Length 1353;
 Best Local Similarity 85.7%; Pred. No. 7.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVG 7
 Db 514 ELYENVG 520
 RESULT 11
 ID Q88GB4 PRELIMINARY; PRT; 133 AA.
 AC Q88GB4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Conserved hypothetical protein.
 GN Pp3810.
 OS *Pseudomonas putida* (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moarzez A.,
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hchisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmeler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile *Pseudomonas putida* KT2440."
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AE016788; AAN69404.1; -;
 DR TIGR; PP3810; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 133 AA; 15335 MW; 82D75532F236679 CRC64;
 Query Match 82.1%; Score 32; DB 16; Length 133;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVGMX 9
 Db 63 RLYENVLGR 71
 RESULT 12

Q9SV79 ID Q9SV79 PRELIMINARY; PRT; 231 AA.
 AC Q9SV79;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN F25G13.3 OR At4G12900.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1] _
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Lemcke K., Schueller C.,
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-13 FROM N.A.
 RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Robben J., Grymoprez B., Volckaert G., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL079349; CAB53090.1; -;
 DR EMBL; AL161535; CAB78332.1; -;
 DR FIR; H85138; H85138.
 DR GO; GO:0004182; F:carboxypeptidase A activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR004911; GILT.
 DR InterPro; IPR000834; Peptidase_M14.
 DR Pfam; PF03227; GILT; 1.
 DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 231 AA; 26025 MW; 734109A78E942295 CRC64;
 Query Match 82.1%; Score 32; DB 10; Length 231;
 Best Local Similarity 71.4%; Pred. No. 1.9e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVG 7
 Db 182 PLYENIG 188
 RESULT 13
 Q8SML6 ID Q8SML6 PRELIMINARY; PRT; 283 AA.
 AC Q8SML6;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Chloroplast large-subunit ribosomal RNA (rrnL), site-specific DNA
 DE endonuclease I-DpaI genes.
 OS Dunaliella parva.
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Dunaliellaceae; Dunaliella.
 OX NCBI_TaxID=3048;
 RN [1] _
 RP SEQUENCE FROM N.A.

RA Turnell M., Otis C., Mercier J.-P., Guttell R.R., Lemieux C.;
 RT "Distribution of group I introns in the chloroplast large subunit rRNA
 gene of green algae";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L43540; AAL77562.1; -;
 DR GO; GO:0009507; C:chloroplast; IEA.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR InterPro; IPR004860; LAGLIDADG_2.
 DR Pfam; PF03161; LAGLIDADG_2; 1.
 KW Chloroplast.
 SQ SEQUENCE 283 AA; 32080 MW; 2BD710EC7BEC1E82 CRC64;
 Query Match 82.1%; Score 32; DB 8; Length 283;
 Best Local Similarity 55.6%; Pred. No. 2.4e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XLYENVGMX 9
 Db 197 ALYENLGIE 205
 RESULT 14
 Q9X169 ID Q9X169 PRELIMINARY; PRT; 352 AA.
 AC Q9X169;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein TM1348.
 GN TM1348.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1] _
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 Haft D.H., Hickley E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linner K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 genome sequence of Thermotoga maritima";
 RL Nature 399:323-329(1999).
 DR EMBL; AB001789; AAD36419.1; -;
 DR FIR; D72264; D72264.
 DR TIGR; TM1348; -;
 DR InterPro; IPR001440; TPR.
 DR InterPro; IPR008941; TPR-like.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 352 AA; 41237 MW; 47EF0B432D421CB8 CRC64;
 Query Match 82.1%; Score 32; DB 16; Length 352;
 Best Local Similarity 55.6%; Pred. No. 3e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 XLYENVGMX 9
 Db 330 RLYEEIGMH 338
 RESULT 15
 Q8RBB7 ID Q8RBB7 PRELIMINARY; PRT; 365 AA.
 AC Q8RBB7;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Bacterial cell division membrane protein.
 GN FTSW OR TTR0905.
 OS Thermoanaerobacter tengcongensis.

OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.

OX NCBI_taxID=119072;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MB4 / JCM 11007;

RX MEDLINE=2192816; PubMed=11997336;

RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,

RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,

RA Tan H., Chen R., Wang J., Yu J., Yang H.;

RT "A complete sequence of *T. tengcongensis* genome."

RL Genome Res. 12:689-700(2002).

DR EMBL; AB013057; AM24161.1; -.

DR GO; GO:000910; P:cytokinesis; IEA.

DR InterPro; IPR001182; Cell cycle.

DR Pfam; PF01098; FTSW RODA_SPOVE; 1.

DR PROSITE; PS00428; FTSW RODA_SPOVE; 1.

KW Cell division; Complete proteome.

SQ SEQUENCE 365 AA; 40320 MW; OCCAEC254B1E81E2 CRC64;

Query Match

Best Local Similarity 82.1%; Score 32; DB 16; Length 365;

Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 XLYENVGMX 9

DB :::|::|:

314 HIFENIGMT 322

Search completed: July 15, 2004, 07:30:42

Job time : 35 secs

